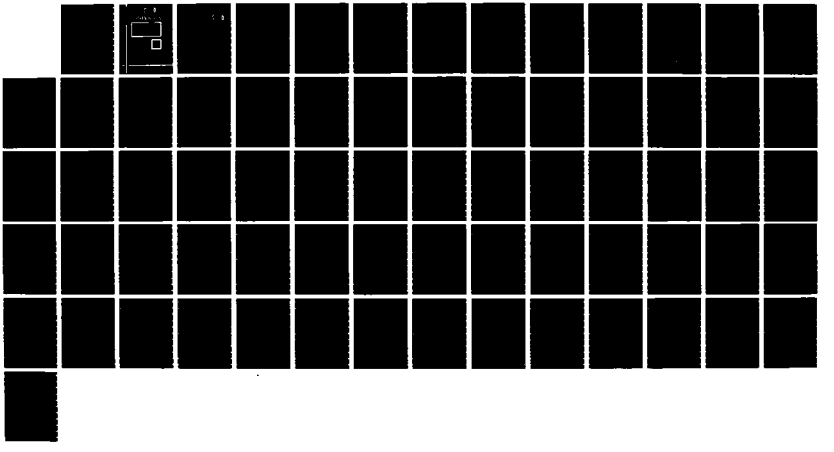
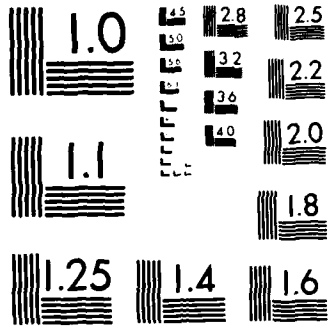
AD-A171 574 AN ASSESSMENT OF SEMICONDUCTOR ALLOYS: PHYSICS AND APPLICATION(U) AUBURN UNIV ALA DEPT OF PHYSICS A CHEN 08 JUL 86 N00014-84-K-0546 1/1

UNCLASSIFIED F/G 20/12 NL





MICROCOPY RESOLUTION TEST CHART

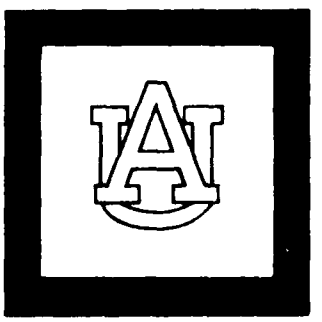
PHYSICS

AN ASSESSMENT OF SEMICONDUCTOR ALLOYS:
PHYSICS AND APPLICATION

Final Report Submitted to
Office of Naval Research

Contract No. N00014-84-K-0546

July 8, 1986



AUBURN UNIVERSITY
AUBURN, ALABAMA

AN ASSESSMENT OF SEMICONDUCTOR ALLOYS: PHYSICS AND APPLICATION

Final Report Submitted to
Office of Naval Research

Contract No. N00014-84-K-0546

July 8, 1986

Submitted by

An-Ban Chen
Physics Department
Auburn Univeristy
AL 36849

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AASA02	2. GOVT ACCESSION NO. ADA171 574	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Assessment of Semiconductor Alloys: Physics and Application		5. TYPE OF REPORT & PERIOD COVERED Final 84 August 1 - 86 June 30
		6. PERFORMING ORG. REPORT NUMBER AASA02
7. AUTHOR(s) An-Ban Chen		8. CONTRACT OR GRANT NUMBER(s) N00014-84-K-0546
9. PERFORMING ORGANIZATION NAME AND ADDRESS Auburn University Physics Department Auburn University, AL 36849-3501		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 410
11. CONTROLLING OFFICE NAME AND ADDRESS Scientific Officer, Mathematical Science Div. Office of Naval Research, 800 N. Quinch St. Arlington, VA 22217		12. REPORT DATE 86 July 8
		13. NUMBER OF PAGES 4
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

The United State Government is authorized to reproduce and distribute this report for governmental purpose

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

monograph, semiconductor alloys, structure and bonding, statistics, phase diagram, alloy scattering, defect, mechanical properties, superlattice

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See Reverse Side

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

This contract supported my work toward a monograph on semiconductor alloys. The activities supported included a one-year sabbatical leave to Stanford University, the collection of information for the book, and the development of theoretical interpretations. All these activities have been accomplished, and the monograph is expected to be completed in 1986. The chapters to be included are: Crystal structure and bonding, Alloy statistics and phase diagram, Electronic structure, Mechanical properties and lattice vibration, Defect states and structures, Alloy scattering effects, and Material engineering -- alloy and superlattice. A copy of the monograph, when published, will be forwarded to ONR to acknowledge the support.

ACCOMPLISHMENTS

This contract supported my work toward a monograph on semiconductor alloys. This a very fast growing area of research. Although we have collected an extensive data base for the book, new data are pouring in and are changing every day. Theoretical work has also been increasing substantially in past two years; yet the interpretations for many alloy topics are still very rudimentary. Among the most striking discoveries are the bi-modal bond length distributions in pseudo-binary alloys, the long-range ordering of epitaxially grown alloys such as GaAlAs, and new results for the band offsets. Against this background, we had to modify our work constantly in order to prevent the book from becoming obsolete. As a result, we have further extended the deadline of the book to the end of 1986. However, all the activities supported by this contract have been accomplished, which included a one-year sabbatical visit to Stanford University, the collection of the data base for the book, and the development of theoretical interpretations. Instead of further extending the expiration date of the contract to coincide with the deadline of the book, I should conclude the contract now. A copy of this monograph, when published, will be forwarded to ONR to acknowledge the support. Below is an outline of the contents:

Chapter One: Crystal Structure and Bonding

1. Crystal structures of important semiconductors, Tables: Lattice parameters, Atomic term values and covalent radii.

2. Tetrahedral bonding: covalent and ionic nature, Harrison's model and modified version, first-principles theory, Table: experimental and theoretical cohesive energies.

3. Structure of pseudo-binary alloys revealed from EXAFS and theoretical models, Table: Alloy bond lengths.

4. Other alloy structures, clustering, and long-range order.

Chapter Two: Alloy Statistics and Phase Diagram

1. Mixing free energy, enthalpy and entropy, miscibility gap, spinodal decomposition, critical temperature, Table: Thermodynamics data for binary compounds.

2. Statistical models: ideal alloy, zeroth approximation, quasi-chemical approximation (QCA), cluster variational method (CVM). Tables: Comparison of theoretical T_c.

3. Models for mixing enthalpies, Table: Estimated and measured values.

4. Phase diagram, Vieland method, Binary liquidus, Ternary liquidus-solidus curves.

5. Generalized QCA.

6. The 16-bond clusters and non-random distributions.

Chapter Three: Electronic Structure

1. Elements of pure-crystal band structure.

3. Alloys as means for tailoring the band gaps.

4. Bowing of alloy band gaps, Table: experimental and theoretical bowing parameters.

5. The coherent potential approximation and the realistic tight-binding theory.

6. CPA band structures for Si_xGe_{1-x}, and III-V and II-VI pseudo-binary alloys.

7. Effects of non-random distribution.

8. Magnetic semiconductor alloys.

Chapter Four: Mechanical Properties and Lattice Vibration

1. Elasticity of pure semiconductors, Table: Bulk modulus and stiff coefficients.

2. Elastic models.

3. Hardness, Table: Measured hardness of semiconductors.

4. Pure semiconductor vibrational bands, Table: LO and TO modes.

5. Defect modes, Table: Local modes from IR and Ramann spectra.

6. Alloy lattice vibration, mixed and separate modes, Table: Experimental alloy LO and TO modes.

7. Analysis using the statistical cluster model.

Chapter Five: Defect states and structures

1. Simple defect states, Table: Identified shallow and deep levels of simple defects in pure semiconductors.

2. Defect levels in alloys.

3. Defect formation energy and diffusion barrier, Table: Summary of theoretical and experimental results.

4. A comparative study of vacancy formation energies in III-V and II-VI compounds and alloys using bond-orbital model.

Chapter Six: Alloy Scattering Effects

1. Electron and hole mobilities and alloy scattering.

2. Brooks' formula and the generalized version.

3. Survey of low-field mobilities in alloys.

4. Velocity-field characteristics in semiconductors and alloys.

5. Disorder effects on optical and photoemission spectra.

Chapter Seven: Material Engineering -- alloy and superlattice

1. Band-gap engineering: examples.

2. Band alignments, Table: Valence-band offsets and Schottky Barriers.

3. Band gaps in superlattices, effects of strain and inter-layer diffusion, quantum-well models.

4. Tight-binding theory for surface, interface, and superlattice.

5. Structural improvement by alloying

6. Future expectation

COLLABORATIONS

Dr. Arden Sher at SRI International is the coauthor of the book. Collaborations with him and his group, including Drs. Krishnamurthy, Berding and Van Schilfgaarde, are indispensable. I also benefitted from interactions with Dr. W. E. Spicer and his group, particularly during my sabbatical leave at Stanford University. These interactions have resulted in a number of publications listed below. These results constitute an important part of the theoretical interpretations used in the book. The interaction with Dr. R. Graft at the Army Night Vision Laboratory is also very useful in the development of the superlattice theory.

PUBLICATIONS

1. "Effects influencing the structural integrity of semiconductors and their alloys", A. Sher, A.-B. Chen , W. E. Spicer, and C.-K. Shih, J. Vac. Sci. Techno. A3(1), 105 (1985)

2. "Dislocation energies and hardness of semiconductors", A. Sher and A.-B. Chen, Appl. Phys. Lett. 46 (1), 54 (1985)

3. "Phenomena influencing the dislocation density of semiconductors and alloys", A. Sher, A.-B. Chen, and W. E. Spicer, Thirteen International Conference on Defects in Semiconductors, edited by Limering and Parsey, Jr. (The Metallurgical Soc. of AIME), 335 (1985)

4. "Sensitivity of defect energy levels to host band structures and impurity potentials in CdTe", A.-B. Chen and A. Sher, Phys. Rev. B31, 6490 (1985)

5. "Binding energy and spectral width of Si 2p core excitons in Si_xGe_{1-x} alloys", S. Krishnamurthy, A. Sher and A.-B. Chen, Phys. Rev. Lett. 55, 320 (1985)

6. "Generalized Brooks' formula and the electron mobility in Si_xGe_{1-x} alloys", S. Krishnamurthy, A. Sher and A.-B. Chen, Appl. Phys. Lett. 55, 320 (1985)

7. "Semiconductor pseudo-binary alloys: bond-length relaxation and mixing enthalpies", A.-B. Chen and A. Sher, Phys. Rev. B32, 3695 (1985)

8. "Band structure of Si_xGe_{1-x} alloys", S. Krishnamurthy, A. Sher and A.-B. Chen, Phys. Rev. B33, 1026 (1986)

9. "Correlations in pseudo binary alloys", A. Sher, A.-B. Chen and M. Van Schilfgaarde, accepted for publication in J. of Vac. Sci. Techno. (1 6)

10. "Coulomb energy in pseudo binary alloys", M. Van Schilfgaarde, A.-B. Chen and A. Sher, submitted to Phys. Rev. Lett. (1986)

11. "Velocity field characteristics of III-V semiconductor alloys", S. Krishnamurthy, A. Sher and A. B. Chen, submitted to J. of Appl. Phys. (1986)

REPRINTS OF PUBLICATIONS

PHENOMENA INFLUENCING THE DISLOCATION DENSITY OF SEMICONDUCTOR COMPOUNDS AND ALLOYS [a]

by

A. SHER, AN-BAN CHEN and W.E. SPICER

Reprinted from:

Thirteenth International Conference on Defects in Semiconductors

Proceedings of the 13th International Conference on Defects in Semiconductors sponsored by the Electronic Materials Committee of The Metallurgical Society of AIME, held at Coronado, California, August 12-17, 1984.

Edited by

L. C. Kimerling
AT&T Bell Laboratories
Murray Hill, New Jersey
and
J. M. Parsey, Jr.
AT&T Bell Laboratories
Murray Hill, New Jersey

CONFERENCE PROCEEDINGS

The Metallurgical Society of AIME

PHENOMENA INFLUENCING THE DISLOCATION DENSITY OF SEMICONDUCTOR COMPOUNDS AND ALLOYS[a]

A. Sher, SRI International, Menlo Park, CA 94062
An-Ban Chen, Auburn University, Auburn, AL 36489
W.E. Spicer[b], Stanford University, Stanford, CA 94305

Abstract

The objective of this paper is to identify the principal microscopic phenomena controlling dislocation densities in bulk grown semiconductors. Then, based on this understanding, a strategy for selecting materials to reduce dislocation densities is offered. The relevant quantities are calculated from an extension of Harrison's bonding theory, which, with our improved accuracy relates properties of the solids to the constituent atoms' valence electron energy states and wave functions. We report on the alloy composition variation of bond energies, bond lengths, charge redistribution among constituents, vacancy formation energies, dislocation energies, and hardness. Several III-V and II-VI compound semiconductors are treated including, GaAs, GaInAs, HgCdTe, and ZnHgTe.

(a) This work was supported in part by AFOSR Contract F49620-81-K0012 and DARPA Contract MDA-903-83-C-0108.

(b) Stanford Aschermann Professor of Engineering

Introduction

According to a currently accepted model (1) of the mechanism leading to dislocations in bulk-grown material at a given temperature and temperature gradient behind the growth front; their density is governed by:

- The vacancy density that depends on the formation energies,
- The competition between vacancy annealing rates and vacancy interaction caused clusters,
- The condensation rate of these vacancy clusters into dislocation loops, and
- The subsequent growth rate of these loops.

The objective of this paper is to identify some of the principal microscopic phenomena controlling these features in semiconductors compounds and their pseudobinary alloys. If any of the foregoing steps can be inhibited, then there will be fewer dislocations in bulk-grown crystals.

A model of the bonding of tetrahedrally coordinated semiconductor compounds due to Harrison underlies this work (2). He has derived expressions for the bond energies and strain coefficients of pure semiconductor compounds in terms of the constituent atom's valence state wave functions and energies. There are four contributions to the bond energies:

- A covalent energy arising from the interaction between sp^3 hybrids on adjacent sites, which, according to a universal rule deduced by Harrison, is related to the inverse square of the bond length, d,
- An ionic energy which is proportional to the energy difference between the sp^3 hybrid energies of the anion and cation,
- A metallic energy arising from the interaction between filled bonding and unfilled antibonding states on adjacent bonds, and
- An overlap repulsion energy which is taken to vary as d^{-4}.

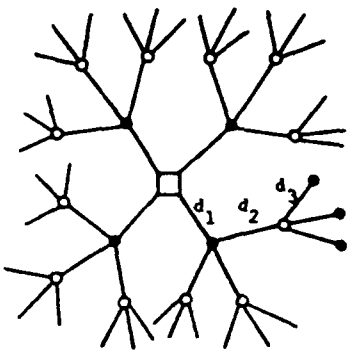
The shear coefficients, which play an important role in dislocation energies, are shown to vary for covalent solids as the covalent energy per cell unit volume or as d^{-5}. The ionic energy arising as it does from coulomb interactions, is insensitive to bond angle distortions and depends only on bond lengths. Thus, in the ionic semiconductors the shear coefficients are reduced from those of equivalent bond length covalent compounds, but also the power law dependence on bond length increases to d^{-11} for the extreme ionic limit (3). The metallic interaction is bond angle sensitive and contributes to the d^{-11} power law.

We have modified Harrison's theory in several ways (3) (4). First, instead of using atomic term values as the input energies in the bonding calculation, we use correlated atomic energies that take account of all the atomic electronic energy changes associated with a state change. In our prior publications these energies were calculated from pseudopotentials (3). In this paper we have modified the procedure and taken the ground state energies from free atom ionization experiments, and only calculate the excitation energies from the pseudopotentials. Second, the Harrison's theory has been extended to alloys in which each type of the individual bond's energy and length changes and the net alloy substitution energies are calculated. A theory of dislocation energies and hardness of semiconductors has been advanced.(4) In this paper the second item will be emphasized since item one is less relevant to this conference and these will be published elsewhere.

Bond Length and Energy Changes in an Alloy

Single Impurity

Start by considering the simplest defect, one cation of a host semiconductor compound AC is replaced by an isoelectronic impurity I, as sketched in Fig. 1. This will be generalized to an $AB_{1-x}C_x$ alloy later ans the cation substituted case $A_{1-x}B_xC$ follows by symmetry. To study the principal effect of core lattice distortions, we allow the positions of the first shell of atoms labeled by bond length d_1 to move, but fix the atoms of the second neighbors and beyond at their pure-crystal positions. Eventually, the effects of long range strain fields, generated by a point distortion, must be added to this core distortion. Here we allow only three different bond lengths, $d_1 = (1-\delta)d$, $d_2 = (1 + 2/3\delta + \delta^2)^{1/2}d$, and $d_3 = d$, where d is the bond length of the pure host AC compound, and δ is a scale parameter.



Schematic representation of the bonds around an atom of interest (designated by the square). The first, second, and third neighbor bonds are designated d_1, d_2, and d_3

There are four bonds with energy E_1, twelve with energy E_2, and 36 with energy E_3. These energies can be written in closed form in the Harrison theory. The energies E_1 and E_2 differ from the host bond energy E_o because of bond length and angle distortions as well as difference in the ionic energy of I and C atom induced charge transfer. The energy E_3 differs from E_o only by the charge transfer coupled through the metallization terms. There are two interesting energies we can calculate. The first is the energy Δ_s required to substitute the I atom for a C atom, i.e. bring a free I atom from infinity and replace a C atom that is taken from the crystal and removed to infinity.

$$\Delta_s = 4\Delta E_1 + 12\Delta E_2 + 36\Delta E_3 - (\epsilon_C - \epsilon_I) \qquad (1)$$

where $\Delta E_j = E_j - E_o$, $j = 1,2,3$, and ϵ_C and ϵ_I are the free atom valence electron binding energies. The second is the bond energy change of the impurity AI bond relative to the bond energy of a pure AI compound, denoted Δ_I

$$\Delta_I = \Delta E_1 - \frac{1}{4}(\epsilon_I - \epsilon_C) \qquad (2)$$

This energy tells us if the AI bond is stabilized ($\Delta_I < 0$) or destabilized ($\Delta_I > 0$) when it is in an AC host. If we define Δ_b as the bond energy difference between the IA and CA bonds each in their respective pure crystals

$$\Delta_b = BE(IA) - BE(CA) \qquad (3)$$

then one can write

$$\Delta_s = 4\Delta_b + \Delta_{sc} \qquad (4)$$

where Δ_{sc} is extra energy difference caused by strain and charge transfer.

The energyes E_j and bond lengths d_j, $j = 1, 2, 3$, are calculated by minimizing Δ_{sc}. If we let $d_I = d_o (1 - \delta_o)$ where d_I is the bond length in the pure AI lattice, then the approximate expression

$$\delta \approx \frac{\delta_o}{1 + \frac{1}{3}\frac{B}{B_I} + \frac{8(C_{11} - C_{12})}{27\, B_I} + (\text{small terms due to charge transfer})} \quad (5)$$

can be deduced. The bulk modulii B of the host and B_I of the impurity lattices, and a shear coefficient $C_{11} - C_{12}$ of the host appear in the expression. While this expression is approximate, the effects of the various types of strains can be visualized using it. Our detailed calculations are done using the full theory. If one fixes the bond lengths $d_1 = d_2 = d_3 = d_o$, then $\delta = \delta_o$ and the strain energy Δ_{sc} is large and positive. If one allows only bond length distortions ($C_{11} - C_{12} = 0$) then δ is reduced and the bonds tend to adjust so $d_1 \approx d_I$. This reduces Δ_{sc} considerably from the undistorted lattice case. If $d_I > d_o$ then both bonds are stretched somewhat, $d_1 - d_I > 0$ and $d_2 - d_o > 0$. However, this configuration produces large bond angle distortions. When the shear coefficients are turned on the lattice relaxation is modified and the bond lengths cannot adjust as much, so d_1 differs from d_I by a larger amount. The net result is that Δ_{sc} is increased since one pays the price of strain energy either as a bond length or bond angle distortion.

If we calculate δ from the full theory, then we predict values that are too large. This occurs because the theory predicts $C_{11} - C_{12}$ which agree well with experiment but it predicts bulk modulii with the proper trends from one compound to another but which are about a factor of 2.4 too small. If we use the experimental values of the strain coefficients in Eq. (5) then good agreement is found with the experiments (5) on $Ga_{1-x}In_xAs$ and $ZnSe_{1-x}Te_x$. The results are quoted in Table I.

Table I. Bond Lengths in Å for Impurities in the Indicated Host

	Ga in InAs	In in GaAs	Se in ZnTe	Te in ZnSe
Experiment[5]	2.487	2.587	2.496	2.595
Eq.(5) and experimental B&C	2.499	2.547	2.510	2.570
Full Theory	2.538	2.518	2.570	2.512

Alloy

In an $A_{1-x}B_xC$ alloy the four cations around a given C anion can be arranged in five different configurations, denoted by $A(4-\eta)$, $B(\eta)$, $\eta = 0,1,2,3,4$. An A(2) B(2) configuration, for example, is one in which the C atom has two A atom and two B atom neighbors. It is possible to solve the full alloy problem for large clusters, but for now we have restricted the cluster to near neighbors only, and in Fig. 1 the central atom is now an anion C and $d_2 = d_3 = d_{eff}$ are taken to be an effective medium bond length that is determined self-consistently. The different bonds of type 1 no longer have the same length. The procedure is as follows: First assume a value for d_{eff}, say the virtual crystal bond length $(1-x)\ d_o(AC) + x d_o(BC)$. Next calculate the d_1 values for the various configurations by minimizing the energy of the configuration. Then, configuration average the various d values to find a new d_{eff}. Then iterate this procedure until it converges.

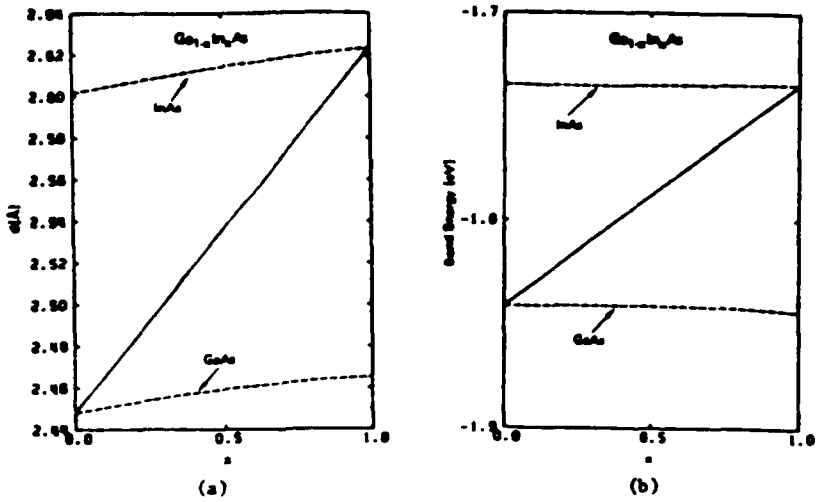
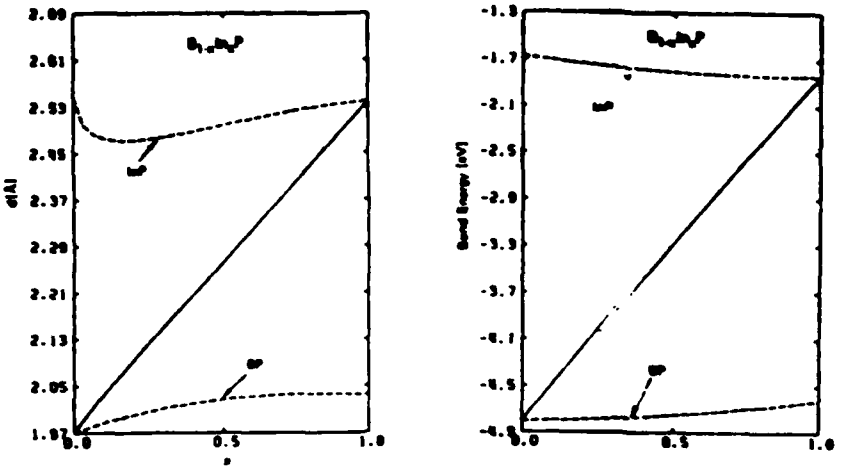


Figure 2: Bond length (a) and energy (b) of $Ga_{1-x}In_xAs$ as functions of x. The dashed curves are the average values for the designated bonds, and the solid curved are the alloy averages.

Because the bond angle distortion terms are unphysically large we have left them out of the present calculation. Hence bond length shifts will be somewhat too small.

Results for the bond energy and bond length shifts with concentration are given for $Ga_{1-x}In_xAs$ in Fig. 2 and $B_{1-x}In_xP$ in Fig. 3 as examples of two different behavior patterns. The predicted trends for GaInAs agree with experiment but the bond length changes are too small as expected with the bond angle distortion terms absent. Notice that the longer bond in this case decreases in length and the shorter one increases as expected. However, in the BInP case the charge shift terms are so large that the longer InP bond has a minimum as a fucntion of composition rather than a monotonic decrease. Because of the large bond length difference between BP and InP there is probably a missability gap in this alloy that prevents these compositions from being prepared. However, one may be able to examine the variation of the anomalous behavior of the InP bond in a BP host ($x \ll 1$).

The $Hg_{1-x}Cd_xTe$ system is completely anamolous. The bond lengths of CdTe (2.805 Å) and HgTe (2.797 Å) are nearly the same by an accident. CdTe bonds are dominated by covalent and ionic interactions, while HgTe is more covalent and the metallic terms are important. The mix of interactions in the two cases leads to the same bond lengths. When an alloy is made the charge shift terms dominate, and they cause the longer CdTe bond to become still longer and the HgTe to contract by amounts that are large compared to the pure crystal difference. Moreover, the already weak HgTe has its bond energy reduced still more by the presence of Cd. Since the melting point of HgTe increases as Cd is added and the local strength of HgTe bonds adjacent to a Cd decrease, vacancy densities will increase. All these and other observed phenomena are predicted by the theory. A complete catalog of results requires more pages than we have been allocated in this article, but we have tried to offer a representative group.



(a) (b)

Figure 3: Bond length (a) and energy (b) of $B_{1-x}In_xP$ as functions of x. The dashed curves are the average values for the designated bonds and the solid curves are the alloy averages.

Conclusions

We have demonstrated in agreement with experiment that dislocation energies and hardness of semiconductors are proportional to the shear coefficients which vary roughly as d^{-9}. Thus to decrease dislocations in a given semiconductor one should inhibit their formation by introducing some means of shortening bonds. This should be accomplished without also decreasing the ratio of the vacancy formation energy to the melt temperature. While detailed calculations to support the following contentions are still incomplete, a set of criteria on an impurity (denoted I) in a host semiconductor (denoted H) that are likely to satisfy these conditions are: the bond length of the impurity is smaller than that of the host $d_I < d_H$, the smaller bond energy is stabilized in the alloy $\min\{\Delta E_{bH}, \Delta E_{bI}\} < 0$. In these circumstances, the average bond length will shrink and in the vicinity of each impurity the four surrounding bonds nearly have the length of the impurity but are stretched slightly. The next neighboring host bonds are also stretched. The net effect is an arrangement which is more rigid than the unperturbed lattice and consequently the local shear strain energy increases, causing dislocation energies to increase. Examples of this case are $B_{1-x}Ga_xAs$, and $Zn_{1-x}Hg_xTe$. Both B in GaAs (6) and Zn in HgTe (7) have proven to be effective in reducing dislocation densities. Unfortunately, B has a low solubility in GaAs and it is not clear that enough can be gotten in to make it easy to prepare dislocation-free material.

A second case where an improvement occurs is if $d_I > d_H$, and again $\min\{\Delta E_{bH}, \Delta E_{bI}\} > 0$. In this case each of the four longer impurity atom bonds are compressed by the surrounding host bonds and, more importantly, the twelve next neighbor host bonds are also compressed. Once again for small impurity concentrations the net dislocation energy should be increased. However, in this case the effect is competing against a net bond lengthening trend of the alloy which is tending to make it less rigid against a shear, so at higher concentrations the mechanism should cease to function. An example

of this case is $Ga_{1-x}In_xAs$, where In has proven to be effective in dislocation reduction of GaAs (8).

We have demonstrated that our modification of the Harrison bonding theory accruately predicts the observed change of bond lengths in semiconductor alloys and offers guidance to means for dislocation reduction.

References

1. G. Schoeck and W.A. Tiller, Phil. Mag. 5, 43 (1960).
2. W.A. Harrison, Electronic Structure and Properties of Solids, (Freeman, San Francisco, 1980); R.C. Sokel Thesis, Stanford University, 1978; W.A. Harrison Microscience , Vol. 4 (limited distribution SRI International publication 1983) p. 34.
3. A. Sher, An-Ban Chen, and W.E. Spicer "Effects Influencing the Structure Integrity of Semiconductors and Their Alloys", 1984 Workshop on the Physics and Chemistry of Mercury Cadmium Telluride, San Diego, 1984.
4. A. Sher, An-Ban Chen, and W.E. Spicer, "Dislocations Energies and Hardness of Semiconductors", (submitted to Appl. Phys.Lett.).
5. J.C. Mikkelsen, Jr. and J.B. Boyce, Phys. Rev. Lett. 19, 1412 (1982), and private communications.
6. G. Jacob, J. Crystal Growth 59, 669 (1982); Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. 42, 822 (1983).
7. S.L. Bell and S. Sen, IRIS Detector Specialty Group Meeting, Boulder Colorado (1983); T.W. James and B.F. Zuck, ibid.
8. R.N. Thomas, H.M. Hobgood, D.L. Barrett, G.W. Eldridge, M.M. Sopira, and M.C. Driver, "Large Dismeter, Low-Dislocation In Doped GaAs:..." Third Conference on Semi-Insulating III-V Materials, Warm Springs, Oregon 1984.

SEMICONDUCTOR ALLOYS: LOCAL BOND LENGTHS, MIXING ENTHALPIES, AND MICROCLUSTERS

A.B. CHEN* AND A. SHER†
*Physics Department, Auburn University, AL 36849
†SRI International, Menlo Park, CA 94025

ABSTRACT

Several recent theoretical studies of the local structure of semiconductor alloys are summarized. First, dilute limit calculations of local bond lengths and mixing enthalpies are discussed. These calculations include effects due to both bond length and bond-angle distortions, as well as local chemical rearrangements. Then, a new statistical theory of concentrated alloys is described. Deviations from random alloy distributions (microclusters) are predicted.

INTRODUCTION

This paper summarizes our recent theoretical studies directed toward understanding the microscopic structures of pseudo-binary semiconductor alloys $A_xB_{1-x}C$. We first present a detailed calculation of the local bond length relaxation in the dilute limit $x \rightarrow 0$, i.e. the case where an A atom is substituted for a B atom in a BC compound. The mixing enthalpy parameter Ω is found to be related to small excess substitution energies. These excess energies are calculated directly through a minimization procedure. Thus, the accuracy of the predicted Ω is not limited by trying to find small differences between large numbers. The theory is then generalized to concentrated alloys using statistics based on combinations of tetrahedral clusters of five atoms. Our results predict that microclustering occurs in a majority of alloys. We conclude by identifying systematic correlations between the theory and several experiments.

Before discussing the calculation, it is useful to provide some background about the structure of these alloys. It was customary to assume that these alloys have two sublattices in which the C atoms occupy one sublattice, and A and B atoms are randomly distributed on the other. This picture, referred to as the virtual crystal approximation (VCA), implies that the nearest-neighbor (nn) bond lengths in the alloy are the concentration weighed average values, i.e. $d_{AC} = d_{BC} = \bar{d} = x\, d_{AC}^{(0)} + (1-x)\, d_{BC}^{(0)}$ where the values with a superscript (0) denote the pure-crystal values. On the other hand, according to Pauling's covalent radii approximation (CRA), the local bond lengths retain their respective pure-crystal values, i.e. $d_{AC} = d_{AC}^{(0)}$ and $d_{BC} = d_{BC}^{(0)}$

If we define $\delta_0 = (\bar{d} - d_{AC}^{(0)}) / \bar{d}$ and $\delta = (\bar{d} - d_{AC}) / \bar{d}$, then the ratio δ/δ_0 in VCA is zero, but in CRA it is 1. However, Mikkelsen and Boyce[1] found from their EXAFS experiment on $Ga_xIn_{1-x}As$ that the nn bond lengths do not fit either VCA or CRA. Instead, they found the value of δ/δ_0 to be close to 3/4. Since then, similar experiments have been done for a number of zinc-blende pseudo-binary alloys,[1] and the 3/4 rule appears to be quite general.

Mat Res Soc Symp Proc Vol 46 1985 Materials Research Society

DILUTE LIMIT

The dilute limit is the easiest case but is still not trivial. Its solution provides both end-point results ($x = 0$ and 1), as well as insight into the extension to the concentrated alloy case. A complete description of this case is being reported elsewhere; here we summarize the essential results. The substitution energy Δ_s for an A atom replacing a B atom in a BC compound is calculated and minimized to find the relaxed configuration. Δ_s can be written as $\Delta_s = 4\,(\Delta E_b + \Delta E_s + \Delta E_{ch})$, where ΔE_b is the binding energy difference between the AC and BC compounds, ΔE_s is the strain energy, and ΔE_{ch} is a chemical energy shift. All $\Delta E'$ s are energies per bond. Then, $\Delta E = \Delta E_s + \Delta E_{ch}$ is the excess energy per bond for the impurity substitution. ΔE_s is calculated by dividing the crystal into two regions. Outside R (which is the distance of the second-shell atoms to the impurity), the distorted crystal is treated as an elastic continuum with a radial displacement field which is inversely proportional to the square of the radial distance, so $\Delta E_s^{(out)} = 1/4RCu^2$, where C is an effective shear coefficient,

$$C = \pi\,(1.6\,(C_{11} - C_{12}) + 4.8\,C_{44}), \tag{1}$$

and u is the magnitude of the displacement at R. Inside R, the strain energy $\Delta_s^{(in)}$ is treated with a valence force field (VFF).[2] Finally, the chemical energy shift ΔE_{ch} is calculated from Harrison's model and arises from changes in the metalization energies caused by different bond lengths $\Delta d \equiv d_{BC} - d_{AC}$ and covalent energies $\Delta V_3 \equiv V_3(AC) - V_3(BC)$. Note that $\delta_0 = (d_{BC}^{(0)} - d_{AC}^{(0)}) / d_{BC}^{(0)}$ and $\delta = (d_{BC}^{(0)} - d_{AC}) / d_{BC}^{(0)}$ in this dilute limit, so the excess energy ΔE can be expanded up to second order in δ, u, and ΔV_3. For a given pair A and B, ΔE is an explicit function of δ and u. Minimization of ΔE with respect to δ and U leads to the equilibrium local bond length d_{AC} and energy ΔE. Then, ΔE is used to estimate the mixing enthalpy parameter Ω in the mixing enthalpy $\Delta H = x(1-x)\Omega$ by

$$\Omega = 2\,(\Delta E\,(A \text{ in } BC) + \Delta E\,(B \text{ in } AC)). \tag{2}$$

A systematic comparison with other models based on strain energy alone shows that an increase of the range of the fixed boundary R increases the relaxation of d_{AC}, i.e. it causes δ/δ_0 to increase. The inclusion of the bond angle restoring force, on the other hand, reduces the relaxation. It turns out that a delicate cancellation of these two effects causes a simple spring model pointed out by Shih et al. (SSHS)[3] to yield accurate results. In this model $\delta = \delta_0 / (1 + 1/3\,\alpha/\alpha_1)$, where α and α_1 are the bond-stretching force constants for the host (BC) and the impurity (AC) crystals. With $\alpha \approx \alpha_1$, this model predicts $\delta = 3/4\delta_0$ for a zinc-blende alloy. Although our full perturbation theory (FPT) and the VFF model of Martin and Zunger (MZ)[4] predict d_{AC} with an average absolute deviation comparable to the experimental uncertainty of 0.01A, the simple spring model is even better.

We note that while in our theory, MZ and SSHS, the Ω values are directly calculated without any adjustable parameters, our theory and SSHS agree with the experiment as well or even slightly better than the one-parameter theories.[5,6] Although our theory predicts a negative Ω value for all three (Ga, Al) alloys, the magnitude ($\Omega = -0.17$ kcal/mole) is too small to account for the ordering of $Ga_xAl_{1-x}As$ grown at 600 to 700°C found recently.[7] The calculated Ω values also provide guidance in separating miscible from immiscible alloys. In a random alloy, the criterion for alloy

mixing for all x is $T > T_c$, where the critical temperature T is given by $T_c = \Omega/2R_g$ with R_g being the universal gas constant. Figure 1 is a plot of T_c/T_2 against $|\delta_0|/|\delta_m|$, where T_2 is the lower of the two constituent's melting temperatures, and $\delta_m = 1.63\chi_m$ with χ_m being the ratio of rms bond length amplitude fluctuation to the bond length at T_2. The simple spring model gives $T_c/T_2 = (\delta_0/\delta_m)^2$ as indicated by the solid curves.

There is an empirical rule stating that a miscibility gap will occur if $|\delta_0|$ between two alloy components exceeds 7.5%. However, if T_c/T_2 is plotted against $|\delta_0|$, the simple spring model would not exhibit a smooth simple quadratic curve, and our theoretical points would be much more scattered. This suggests that $|\delta_0|/|\delta_m| > 1$ is a better criterion than $|\delta_0| > 0.075$. Figure 2 also clearly shows the chemical effects, namely negative and positive chemical energies ΔE_{ch} for cation and anion substitutions respectively. The full theory and the experiments correlate within the experimental uncertainties. The simple SSHS model clearly is an excellent universal representation. However, T_c/T_2 varies faster than quadratically for larger $|\delta_0/\delta_m|$ values, as born out from both the experimental data and the full theory.

CONCENTRATED ALLOYS

Turn now to the concentrated alloy case. First, an improved statistical model is required. We have extended regular solution theory based on pair energies to one for five-atom clusters. For an $A_xB_{1-x}C$ alloy, the building blocks are clusters of A(m)B(4-m)C, where m ranges from 0 to 4. For a given alloy concentration x and for a given set of energies ϵ_m associated with these clusters, we have derived expressions for the cluster population distribution $x_m \equiv \bar{n}_m / N$, where N is the total number of unit cells and $\bar{n}_m$ is the averaged number of cells with A(m)B(4-m)C clusters. The partition function Z is obtained using a steepest descents argument which then yields the mixing Helmholtz free energy ΔF. The result reduces to Guggenheim's tetrahedron case[7] if pair potentials (for the second-neighbors) are assumed. Another major difference is that we only need to solve a single quartic equation, while Guggenheim needed to solve four simultaneous quartic equations.

The key to the problem, however, lies in the calculation of the energies ϵ_m. If one assumes that the size of the tetrahedra for all m-clusters at a given alloy concentration takes on the corresponding VCA values but allows the central C atom to relax, then the energies as functions of x behave like those shown in Figure 2(a). There are at least two major flaws in this result. First, the energies are too large and would correspond to Ω values many times the experimental values. Second, at x = 0.75, 0.5, and 0.2, these energies imply compound formation for $A_3B_1C_4$, $A_2B_2C_4$, and $A_1B_3C_4$, respectively, which is opposite to the known tendency for spinodal decomposition of $Ga_xIn_{1-x}As$ at low T. However, if the local cell volume of each cluster is allowed to be in mechanical equilibrium with a continuous medium with an effective shear coefficient $\bar{C} = xC_{AC}^0 + (1-x)C_{AB}^{(0)}$, where the C value for the pure material is given by Eq. (1), then, the corresponding energies ϵ_m as a function of x are given in Figure 2(b), which now yields a reasonable value of mixing enthalpy and correctly predicts the tendency toward spinodal decomposition at low temperature. With this set of energies, one can then calculate the cluster distribution x_m, and compare them with the corresponding values for a random alloy, i.e. $x_m^{(0)} = (4_m)x^m(1-x)^{4-m}$. Figure 3 shows the deviation from randomness $\Delta x_m = x_m - x_m^{(0)}$ as a function of x for four arbitrarily chosen growth

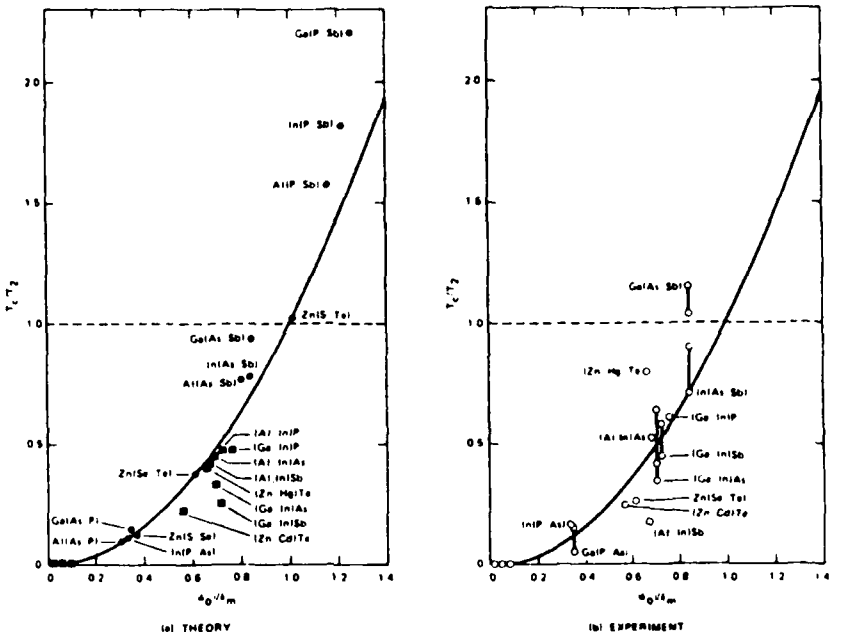


FIGURE 1 PLOT OF T_c/T_2 AS A FUNCTION $|\delta_o - \delta_m|$ FOR T_c OBTAINED FROM (a) PRESENT THEORY AND (b) THE EXPERIMENTAL Ω. IN (a), THE DOTS ARE FOR ANION ALLOYS AND CROSSES FOR CATION SUBSTITUTION. THE SOLID LINES IN BOTH PARTS CORRESPOND TO THE Ω BASED ON THE SIMPLE THEORY OF REF. 3. THE DASHED LINES AT $T_c/T_2 = 0$ SEPARATE THE MISCIBLE FROM IMMISCIBLE GROUPS

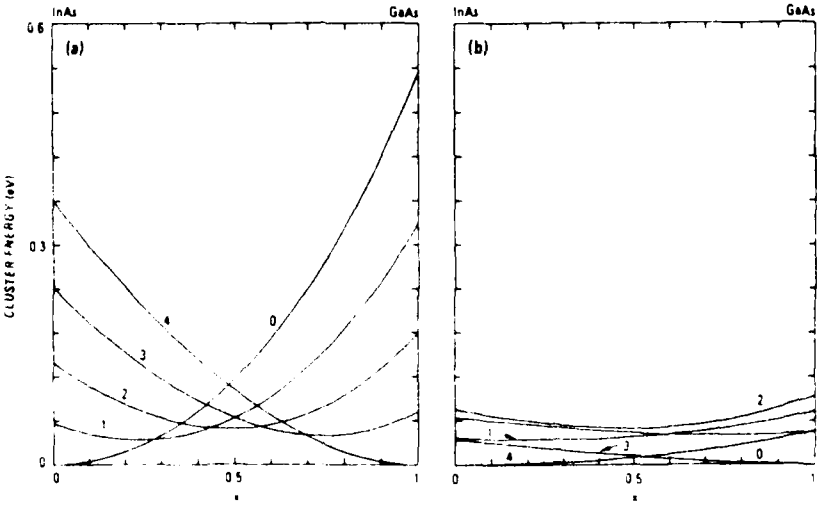


FIGURE 2 THE CLUSTER ENERGIES ϵ_m AS A FUNCTION OF x CALCULATED IN TWO DIFFERENT WAYS, AS DISCUSSED IN THE TEXT

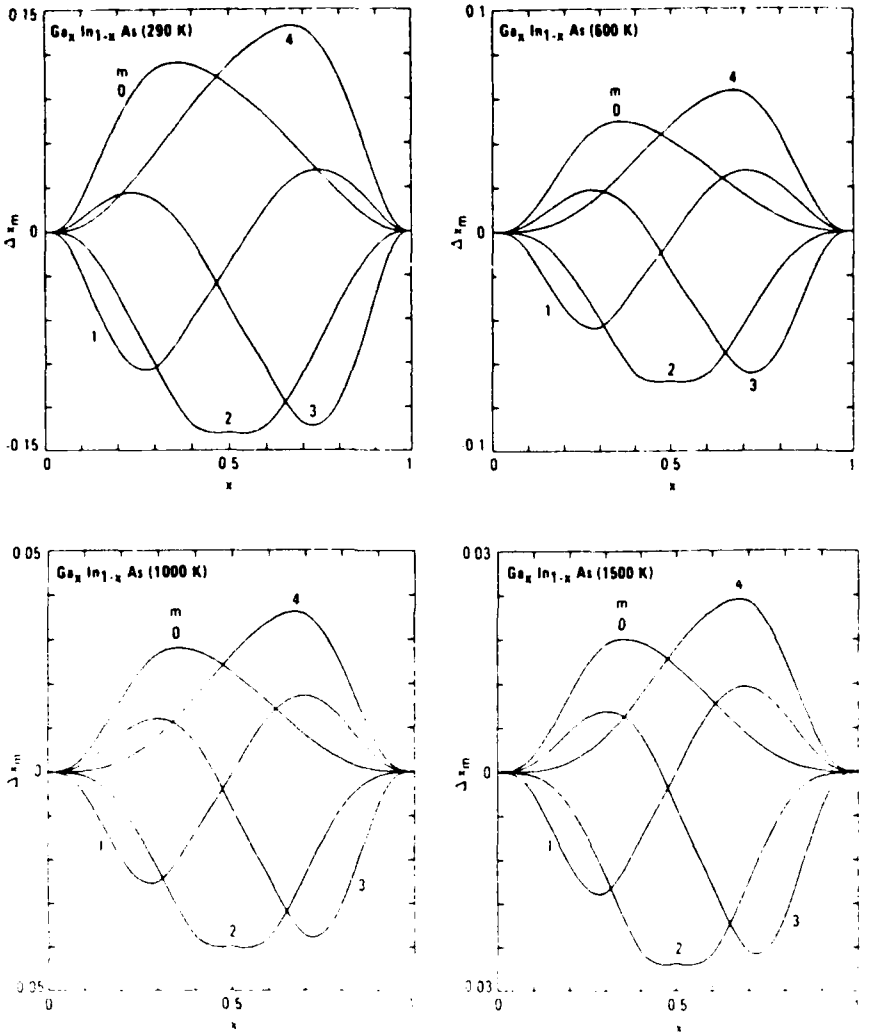


FIGURE 3 DEVIATIONS FROM RANDOM DISTRIBUTIONS FOR THE FIVE CLUSTERS AT SEVERAL GROWTH TEMPERATURES

Dislocation energies and hardness of semiconductors

A. Sher
SRI International, Menlo Park, California 94025

A.-B. Chen
Auburn University, Auburn, Alabama 36849

W. E. Spicer
Stanford University, Stanford, California 95305

(Received 13 July 1984; accepted for publication 5 October 1984)

The dislocation energies and hardness of semiconductors are calculated by an extension of Harrison's method. It is demonstrated in agreement with experiment that dislocation energies per unit length are proportional to d^{-3}–d^{-9}, where d is the bond length and hardness is proportional to d^{-5}–d^{-11}. The hardness is related to the interaction energies among dislocations. It is argued that dislocation densities of semiconductors will be reduced if they are alloyed with a second constituent that has a shorter bond length. Experimental evidence supporting this strategy is noted.

Dislocations in semiconductors are detrimental to device function; they serve as channels for impurity migration and trapping, which cause nonuniform doping and degrades p-n junctions.[1] They also decrease the material's resistance to plastic deformation. The aim of this letter is to provide insights into the underlying physical mechanisms controlling dislocations and semiconductor hardness, and then to suggest strategies for decreasing dislocation densities. It is well established that the hardness of tetrahedrally coordinated semiconductor materials—groups IV, III-V, and II-VI compounds—exhibits a sharp variation with their near-neighbor distance d, approximately proportional to d^{-9} for one group of seven compounds.[2] Thus, semiconductors with small lattice constants tend to be harder materials. These same materials have larger stiffness coefficients[3] and have fewer dislocations in as-grown crystals.[4-6]

The shear coefficients (combinations of C_{44} and $C_{11} - C_{12}$ in the Schoenflies notation)[7] depend on crystal orientation and (in Harrison's notation[8]) are proportional to $V_2^3/d^5(V_2^2 + V_3^2)^{3/2}$, where $V_2 \propto d^{-2}$ is the covalent and V_3 is the ionic energy. The metallic interaction modifies the functional dependence of the shear coefficient on V_2 and V_3, but introduces no explicit dependence on the hopping integrals, denoted V_1 by Harrison.[3,8] In a pure covalent material, the bond energy is proportional to V_2 (or d^{-2}), and the bond volume is $\propto d^3$; hence, in this case, the shear coefficient varies as d^{-5}. In the limit, $V_3 \gg V_2$, $C_{11} - C_{12} \propto d^{-11}$. For most polar semiconductors, d^{-9} is a good approximation.

Hardness is determined by applying a known force F to a probe of a prescribed shape driving it into the surface of the sample.[7] The area A of the resulting indentation is measured, and the hardness is the force per unit indented area. Many dislocations must be formed to allow the probe to indent the semiconductor. If the indenter is a rectangular pyramid, then the hardness is $H = F/A = Fh/Ah = \epsilon_T/Ah$, where ϵ_T is the work required to cause the indenter to penetrate to a depth h. A side view of the indentation in a cut through its center is illustrated schematically in Fig. 1. The top of the indentation has side length W; thus, $A = W^2$. The Burger's vector has magnitude b, proportional to the bond length d. The number of dislocations N_h required to accommodate an indentation to depth h is $N_h = h/b = \frac{1}{2}W \cos\vartheta/b$, where ϑ is the angle between the normal to the tip of the indenter and a side. Figure 1 also shows a model of one possible configuration of the dislocations. The edges of the extra atom planes that are driven from the indented volume into the bulk of the semiconductor are shown as lines terminated by dots. The dotted ends of these lines are the positions of the dislocations, which are perpendicular to the plane of the figure. The planes driven to the sides each have a finite extent and a trapezoidal shape. The planes driven down under the indenter have a square shape.

Much of the work done on the indenter goes into the energy to form the indicated dislocation configuration, although some certainly goes into heat. There are two major contributions to this formation energy. The first is the enegy needed to generate each dislocation as an isolated entity, and the second is the interaction energy among these dislocations. Because the interaction term dominates H, approximations made to simplify the first term are relatively unimportant. The extra planes driven to the sides of the indentation have a finite extent; accordingly (in this idealized picture), there are both edge dislocations at their base and screw dislocations associated with their termination. The square planes driven below the indentation have edge dislocations around the sides and screw dislocations at the corners to make the turns. Moreover, there are interactions

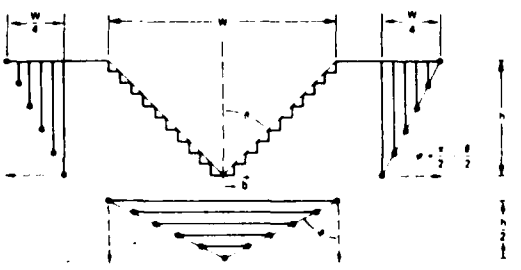


FIG. 1. Schematic representation of an idealized minimum-energy dislocation configuration produced by a square cross-section indenter in a hardness measurement. W is the side length and h is the depth of the indentation. In this ideal case, half the material from the indented region is displaced along the glide planes (indicated by the dashed arrows) to the sides and half is displaced below the indenter.

 0003-6951/85/010054-03$01.00

among the dislocations, which can produce a minimum-energy configuration. For the arrangement depicted in Fig. 1, it always costs energy to position a second dislocation on a parallel glide plane to one already present. However, the magnitude of this extra energy can be minimized and, for proper configurations of the dislocations, there are attractive forces along the glide planes that will tend to position the dislocations into the minimum-energy configuration. The minimum configuration arises when half of the atoms from the indented volume go respectively to the side and below the indenter. Then, in both regions, the maximum angle made between successive close-spaced dislocation lines and their glide planes is $\vartheta = \pi/2 - \vartheta/2$, as shown in the figure. This is the minimum realistic energy configuration. If the dislocations are separated more than shown in Fig. 1, then there is more volume of strained material and the interaction energy would be larger still.

An approximate expression for the energy required to indent the material is[7]

$$\epsilon_I = 2\cdot 4\left[\sum_{i=1}^{N} E_i L_i + \frac{1}{2}\sum_{i,j=1}^{N} E_{ij}\left[\min(L_i, L_j)\right]\right] \tag{1}$$

In the first term, using an isotropic medium approximation and neglecting core terms, the energy per unit length to form an edge dislocation is[7]

$$E_i = \frac{Gb^2}{4\pi(1-\nu)}\ln\frac{R}{r_0}, \tag{2}$$

the shear coefficient is G, the Burger's vector $b = d/3$ for an indentation along a $\langle 100\rangle$ axis, the range of the elastic deformation of a dislocation R is taken equal to W (for want of a better approximation), $r_0 \sim d$ is the dislocation core radius, the Poisson ratio is $\nu \sim 0.2$ for most semiconductors, and L_i is the length of the ith dislocation. In the second term, E_{ij} is the interaction energy per unit length between dislocations i and j. Assuming they have parallel glide planes and their Burger's vectors have the same sign, E_{ij} is given by[7]

$$E_{ij} = \frac{Gb^2}{2\pi(1-\nu)}\left[\ln\left(\frac{R}{r_{ij}}\right) + \cos^2\varphi_{ij}\right], \tag{3}$$

where r_{ij} is the separation, and φ_{ij} is the angle that a line perpendicular to and joining the dislocations makes with the glide plane, as shown in Fig. 1. Because the various dislocations in a region have different lengths, the net interaction energy is approximated by multiplying the energy per unit length by the length of the shorter one. The upper limits on the sums N are the number of dislocations in one region (side or bottom) associated with one edge. For the minimum-energy configuration, $N = N_a/2$. The four that multiplies the bracket accounts for the four sides, and the two for the two regions for each side.

We now encounter our first surprise. As we can see from Eq. (2) and (3), E_i and E_{ij} have comparable magnitudes. Because there are approximately N^2 terms in the interaction energy sum, only N terms in the formation energy sum, and N is a large number ($N \gg 1$), the interaction energy completely dominates the hardness. In fact, N is typically of the order 10^4. Hence, terms owing to screw dislocations, core energies, heat dissipation as the dislocations propagate to their places, and other effects associated with the first term are unimportant. However, care must be taken with the interaction terms: Eq. (1) neglects a number of secondary interactions, some positive and others negative; these will be added later. The principal neglected terms are the interactions between the dislocations in the different regions on each side (positive), and the interactions between adjacent side and bottom regions (positive), and the interaction between the opposite sides in the bottom regions (negative). Comparison of the results with experiment will indicate how important these neglected terms are likely to be. The length L_i is $W(N-i)/N$ for i from 1 to N. This is the largest length of the side-inserted planes and its choice partially accounts for interacitons between the otherwise neglected screw dislocations. The distance $r_{ij} = r_i - r_j$ is given by $2\sqrt{2}b\,|i-j|$ for i and j ranging from 1 to N for the minimum-energy configuration and a tetrahedrally bonded semiconductor. Finally, in the indicated configuration, $\varphi_{ij} = \varphi = \pi/2 - \vartheta/2$. Inserting these expressions into Eq. (1) and retaining only terms of order N^2 yields

$$H_{\min} = \frac{G\cot\vartheta}{6\pi(1-\nu)}\left[-\ln\left(\frac{\cot\vartheta}{\sqrt{2}}\right) + \frac{4}{3} + \sin^2\frac{\vartheta}{2}\right]. \tag{4}$$

One can also get a number for the hardness of a dislocation in which all the material is pushed along the same glide plane, e.g., to the side, to the bottom, or normal to the face of the indenter (a possibility not depicted in Fig. 1). In this case, the factor of 2 in front of Eq. (1) is removed, $N = N_a$, and $\phi = \pi/2 - \vartheta$. Then a higher nonequilibrium hardness in the context of this model (denoted H_1), is obtained

$$H_1 = \frac{G\cot\vartheta}{3\pi(1-\nu)}\left[-\ln\left(\frac{\cot\vartheta}{\sqrt{2}}\right) + \frac{4}{3} + \sin^2\vartheta\right]. \tag{5}$$

The proper answer for most materials, and depending on crystal orientation, probably lies somewhere between $H_{\min}$ and H_1. For an indenter with $\vartheta = \pi/4$, we have $H_{\min} = 0.0969G/(1-\nu)$ and $H_1/H_{\min} = 2.39$. Harrison[8] has shown that one contribution to the shear coefficient (actually $C_{11} - C_{12}$) is $G = 2.38\,\hbar^2 m\alpha_c^5/md^5$, where m is the free-electron mass, α_c is the covalence, $\alpha_c = V_2/(V_2^2 + V_3^2)^{1/2}$, and d is the bond length. We will approximate G by this expression. Using this G and $\nu = 0.2$, and changing the dimensions to those in terms of which experimental hardness numbers are customarily quoted gives $H_{\min} = 2.38 \times 10^4\,(\alpha_c^5/d^5)$ kgm/mm^2, where d is in angstroms. Calculated values of $H_{\min}$ and H_1 are plotted against experimental results in Fig. 2 for a number of semiconductors.

Figure 2 has the theoretical $H_{\min}$ and H_1 values connected by arrows from $H_{\min}$ to H_1 for each compound, plotted as a function of the corresponding experimental values.[7,9] If the theory were perfect and the experimental values were accurate, the points would fall on the indicated unity slope line. Several conclusions can be drawn. Firstly, the order of magnitude of the predicted and measured values are the same, a result obtained with no adjustable parameters in the theory. Secondly, the trends from one compound to another are properly given by the theory. Although the $H_{\min}$ values are generally too small, they fit the soft materials better, and the H_1 values fit the harder materials better. Thirdly, from Eqs. (1) and (3), H is given in a rough but revealing

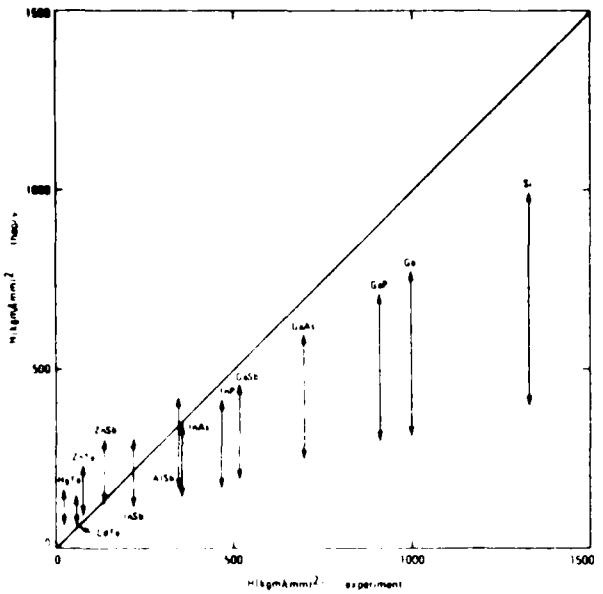


FIG. 2. Theoretical vs experimental hardness of several semiconductors in Refs. 2 and 9. The two theoretical values for each semiconductor are connected by an arrow from H_{min} to H. Perfect agreement would correspond to points being on the unity slope solid line.

approximation by $4N^{-2}E_c l_c/4h$, where the averge dislocation length $l_c = W/2$ and E_c is the average dislocation pair interaction energy per unit length. Notice that N (or W) cancels from this expression; thus, H is independent of W (or F) and therefore, H is truly a measure of the properties of the material. This result would not be found if the dislocation energies ($\propto N$) were to dominate H rather than the pair interaction energy ($\propto N^2$). Finally, the Berger's vector cancels from the leading term and appears only in the argument of the logarithm in Eq. (3). Thus, the answers are also insensitive to its choice.

Dislocations are often found in materials as they are grown. Their density is determined by the thermal and mechanical stresses to which they are subjected in the growth process. A dislocation constitutes a metastable excitation relative to the perfect crystal ground state. At the elevated growth temperatures and temperature gradient behind the growth front, the number of dislocations present is controlled by the relative rate at which vacancies anneal or condense into dislocations.[10] The dislocation formation rate will be slower in a material grown at the same temperature if E_v is higher. If an alloy is formed from a material of interest and a second constituent with a shorter bond length, one expects the average bond energy (and thus both the melt temperature and vacancy formation energy) to increase proportional to a low inverse power of the average bond length.[8] Hence, the equilibrium vacancy density just below the melting point tends to be the same in lowest order for all materials, independent of the bond lengths of the constituents. However, the shear coefficient and dislocation energy per unit length will increase with much higher inverse powers of the bond length. Consequently, dislocation densities should be reduced in such alloys relative to those found in the longer bond length pure constituent. This expectation is confirmed in the recent work reported on $Zn_{1-x}Cd_xTe$ bulk material.[4] The best CdTe that has ever been grown has dislocation densities in excess of 5×10^5 cm^{-2}. The addition of only 4% Zn reduced the dislocation count to less than 5×10^4 cm^{-2}. The ZnTe bond length is 2.643 Å, while the CdTe bond length is 2.805 Å, a 6% difference. This 6% difference in bond length translates into a 2% difference in the dislocation energy per unit length for $1-x=0.04$. Dislocation energies per unit length are typically 10 eV per lattice spacing; accordingly, a 2% increase can be expected to slow their formation rate considerably.

The argument just presented naturally leads to a strategy for decreasing dislocations in other semiconductors. If an alloy is made of the material of interest with another compound with a shorter bond length, then the dislocation density should be reduced. For example, this suggests that the addition of a small amount of GaP ($d=2.359$ Å) may significantly reduce the dislocation density of bulk grown GaAs ($d=2.448$ Å). It has been demonstrated that the addition of approximately 1% GaN ($d=1.946$ Å)[5,6] or of a 10^{18} cm^{-3} BAs ($d=2.069$ Å) concentration[11] to GaAs can yield a large volume of dislocation-free material. An InAs additive with its longer bond length ($d=2.623$ Å) serves the same function indirectly, by causing GaAs bonds in its neighborhood to be compressed. This indirect mechanism should be less effective than substituting short bond length additives.

We have demonstrated that the dislocation energies and hardness of tetrahedrally bonded semiconductors are rapid functions of the reciprocal of the bond length. This rapid d dependence of dislocation energies provides a rationale for the dramatic decrease of the dislocation density in bulk grown $Zn_{0.04}Cd_{0.96}Te$ material relative to that found in CdTe, and suggest means for accomplishing the same ends in other materials.

The authors are indebted to J. P. Hirth, W. A. Harrison, and T. N. Casselman for helpful comments. This work was supported in part by DARPA contract MDA 903-83-C-0108 and AFOSR contract 49620-81-K-0012.

[1] E. M. Swiggard, Proceedings of GaAs IC Symposium, Phoenix, AZ 1983, p. 26.
[2] N. A. Goryunova, A. S. Borshchevskii, and D. N. Fretiakov, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, NY, 1968), Vol. 4, Chap. 1.
[3] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980); R. C. Sokel, thesis, Stanford University, 1978; W. A. Harrison, *Microscience* (limited distribution SRI International publication, Menlo Park, 1983), Vol. 4, p. 34.
[4] S. L. Bell and S. Sen, presented at Infrared Imaging Systems (IRIS) Detector Specialty Group Meeting, Boulder CO, 1983; T. W. James and B. F. Zuck, *ibid*.
[5] Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. **42**, 822 (1983).
[6] G. Jacob, J. Cryst. Growth **59**, 669 (1982).
[7] J. W. Christian, *Theory of Transformations in Metals and Alloys*, 2nd ed. (Pergamon, NY, 1975), Chap. 7. For a discussion of the interaction energies, see J. P. Hirth and J. Lothe, Theory of Dislocations, 2nd ed. (Wiley, NY, 1982), pp. 262–263.
[8] W. A. Harrison, Phys. Rev. B **27**, 3592 (1983).
[9] S. Cole and A. F. W. Willoughby, J. Cryst. Growth **59**, 370 (1982).
[10] G. Schoeck and W. A. Tiller, Philos. Mag. **5**, 43 (1960).
[11] S. Miyazawa, 1983 European Patent Application 83302165.2, filing date 4/18/83, and references therein.

PHYSICAL REVIEW B VOLUME 31, NUMBER 10 15 MAY 1985

Sensitivity of defect energy levels to host band structures and impurity potentials in CdTe

A.-B. Chen
Physics Department, Auburn University, Auburn, Alabama 36849

A. Sher
SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025

(Received 10 December 1984)

The sensitivity of defect energy levels in semiconductors to the host band structures and impurity potentials has been studied for approximately 30 impurities in CdTe using four different band-structure models. The discrepancies in the defect levels between two different sets of band structures and impurity potentials are found to range from less than 0.1 eV to the whole band gap (1.6 eV). The band-structure effects are analyzed here in terms of detailed partial densities of states. Examples of contradictory predictions from different band structures are illustrated, and ways to improve the theory are suggested.

I. INTRODUCTION

In several of our recent papers,[1-7] we have applied a method to calculate the band structure of semiconductors that is both efficient and accurate. Because the procedure involves casting the basis functions into orthonormal local orbitals[6] (OLO), our method has the advantages common to empirical tight-binding (ETB) calculations,[8-10] except that the Hamiltonian matrix elements to all ranges are retained. The inclusion of these higher coefficients makes it possible to produce excellent band structures including conduction bands and effective masses. The method also yields wave functions for optical property calculations.[7] Moreover, its OLO description also permits its extension, through the coherent-potential approximation, to alloys.[2-5]

The recent attention focused on defects in semiconductors motivated us to apply our method to this problem. The theories of defects have ranged from very sophisticated self-consistent density-functional theory[11-13] (SCDF) to simple ETB calculations. It is generally recognized that SCDF is as accurate in defects for the ground-state properties as it is for pure semiconductors, but less certain in assigning excited energy levels. ETB, because it can produce results for many systems in one study, claims to predict the trends of deep levels[10] even if the accuracy for a given impurity may be poor. However, this contention remains to be verified.

To assess this concern, we ask the following question: "How sensitive are defect levels to host band structures and impurity potentials?" To this end, we have adopted the simple yet nontrivial defect model, that of site-diagonal substitutional defects often used in ETB studies. CdTe was selected in this study because its band structure has been examined in great detail by us, and there are three published band-structure models[8-10] that we could easily generate for comparison. There is also a considerable body of experimental data on deep states in this system.[14-17]

II. CALCULATIONAL PROCEDURE

In the simple site-diagonal substitutional defect model, the impurity energy levels E are determined by the equation

$$1-v_\alpha g_\alpha(E)=0\ , \tag{1}$$

where α designates the symmetry of a local state, e.g., Γ_6, Γ_7, and Γ_8 on an atomic site in the zinc-blende structure, and g_α is the real part of the diagonal matrix element of the host-crystal Green function. g_α can be calculated from the partial density of states (PDOS) by

$$g_\alpha(E)=\int \rho_\alpha(\epsilon)/(E-\epsilon)d\epsilon\ . \tag{2}$$

The PDOS is given by

$$\rho_\alpha(\epsilon)=\sum_{n,k} |a_n^\alpha(k)|^2\delta(\epsilon-\epsilon_n(k))\ , \tag{3}$$

where $\epsilon_n(k)$ are band energies and $a_n^\alpha(k)$ are the probability amplitudes of the band state in the Bloch basis constructed from the OLO labeled by α. The Brillouin-zone integration in Eq. (3) is calculated using an accurate ray scheme.[18]

Because a principal concern of this paper is the sensitivity of impurity levels to the host band structures, we should emphasize the difference between our method and ETB. Our method consists of four steps.

(1) We start with four Gaussian orbitals per atom and empirical pseudopotentials,[19] and compute the Hamiltonian matrix $H(k)$ and overlap matrix $S(k)$ as was done by Kane[20] and Chadi.[21]

(2) The Gaussian orbitals are transformed into OLO,[6] so $H(k)$ is transformed into $H_0(k)$ and S into the identity matrix. The band structures calculated from $H_0(k)$ are accurate to 5% as compared to more sophisticated methods using the same potential.[1]

(3) A spin-orbit Hamiltonian in the OLO basis[4] is incorporated to deal with this interaction.

(4) To compensate for the effects of truncated basis and

nonlocal potentials, a perturbation Hamiltonian H_1 is added. H_1 has the same form as a truncated ETB Hamiltonian. The parameters in H_1 are adjusted to fine tune the important band energies and effective masses.[1-4]

Although both ETB and our methods are empirical, there are two major differences.

(1) While most ETB retains the H matrix elements only to the first- or second-neighbor shell, ours extends to all ranges, so that the high Fourier components needed to produce the sharp band curvatures are properly given.

(2) Our method can directly generate wave functions for calculation of other properties.

Thus, while our method yields more accurate band structures, it retains much of the advantage of ETB, namely the computational speed and a simple direct-space description of the Hamiltonian.

III. BAND STRUCTURES AND PARTIAL DENSITIES OF STATES

Figure 1 depicts the four band structures to be considered for CdTe. Our result is in panel (a); panels (b) (Ref. 8) and (c) (Ref. 9) are two ETB band structures with the Hamiltonian matrix elements truncated at second neighbors. (Because different parameters were selected, these two band structures are not identical.) Panel (d) (Ref. 10) results from the use of five basis orbitals per atom; the extra one is an excited s state. All these band structures are adjusted to have the proper fundamental band gap of 1.6 eV. The principal differences one sees on first inspection are in the band curvatures, especially the conduction bands. The effective mass at the bottom of the conduction band in panel (a) is 0.1 times the free-electron mass, in agreement with experiment,[17] while in other panels it is more than twice as large.

Figure 2 shows the densities of states (DOS) for each of the band structures in Fig. 1. While the valence bands at least exhibit general common features, the conduction bands are almost unrecognizable as representing the same compound. In panels (c) and (d), for example, there is a second band gap above the fundamental gap. Also note that there are two extra narrow peaks associated with the two extra excited s orbitals (one for Cd and the other for Te) included in the calculation.

To analyze the band effects on defect levels [see Eqs. (1) and (2)], the DOS is further decomposed into partial densities of states for $\Gamma_6(s)$, $\Gamma_7(p^{1/2})$, and $\Gamma_8(p^{3/2})$ states on the Cd and Te sites, as shown in Figs. 3–6. The Γ_8 PDOS are not shown because they are nearly the same as Γ_7 with only a slight upward energy shift. These PDOS show how the "atomic" levels evolve into band states. These curves contain useful information about many properties, e.g., the relation between the crystal bonding and atomic energies, and how potential disorder in alloys affects different parts of the bands,[2-5] in addition to defect levels studied here.

The Γ_6(Cd) PDOS shown in Fig. 3 split between the conduction and valence bands. It is generally assumed that the cation s states in III-V and II-VI compounds

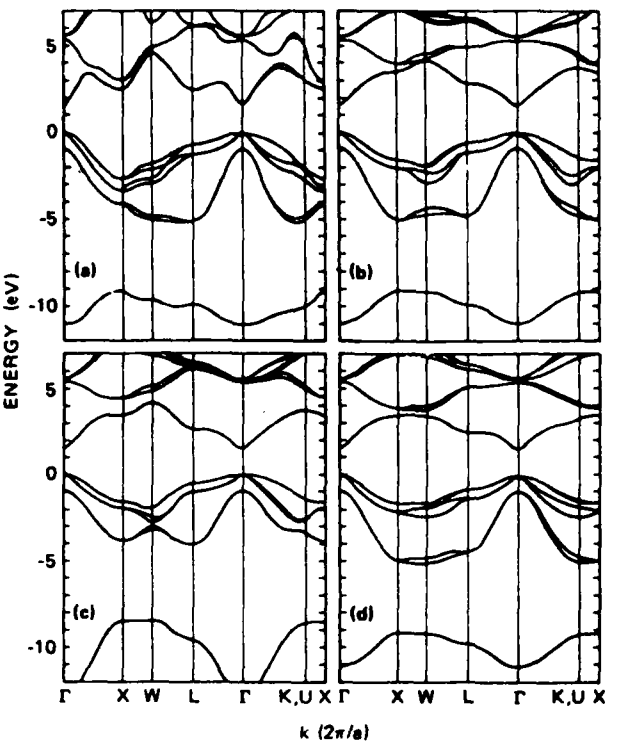


FIG. 1. Four band structures of CdTe used for comparative studies: (a) present work, (b) Ref. 8, (c) Ref. 9, and (d) Ref. 10.

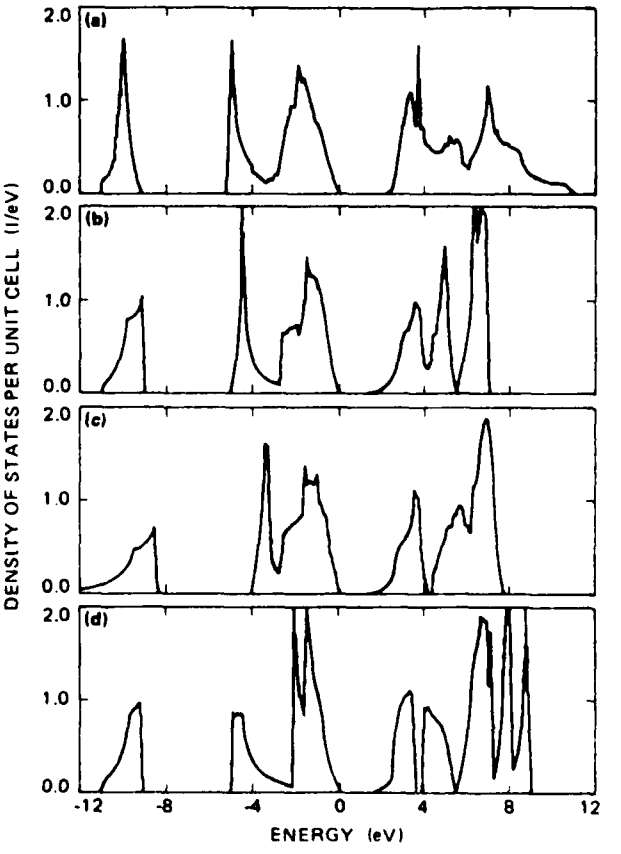


FIG. 2. Densities of states calculated from the four band structures in Fig. 1.

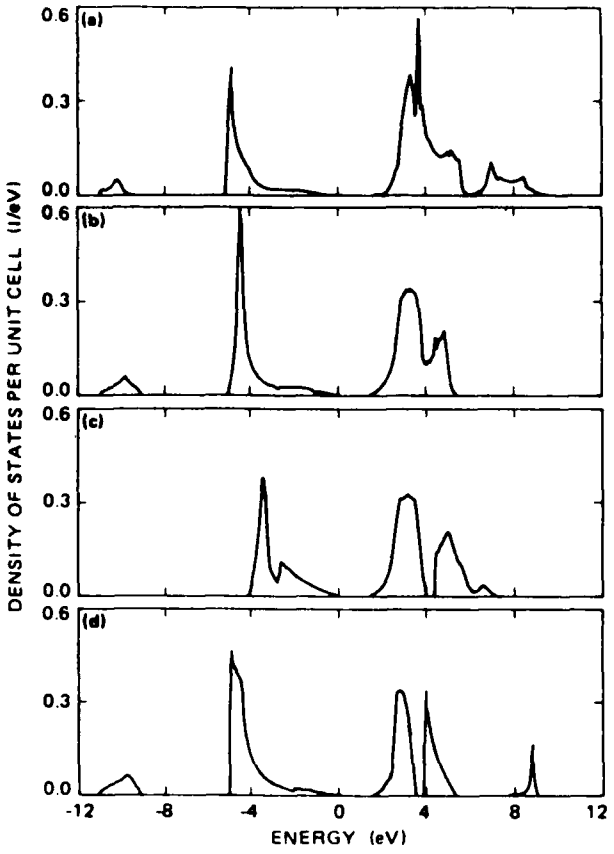


FIG. 3. The Cd Γ_6 partial densities of states.

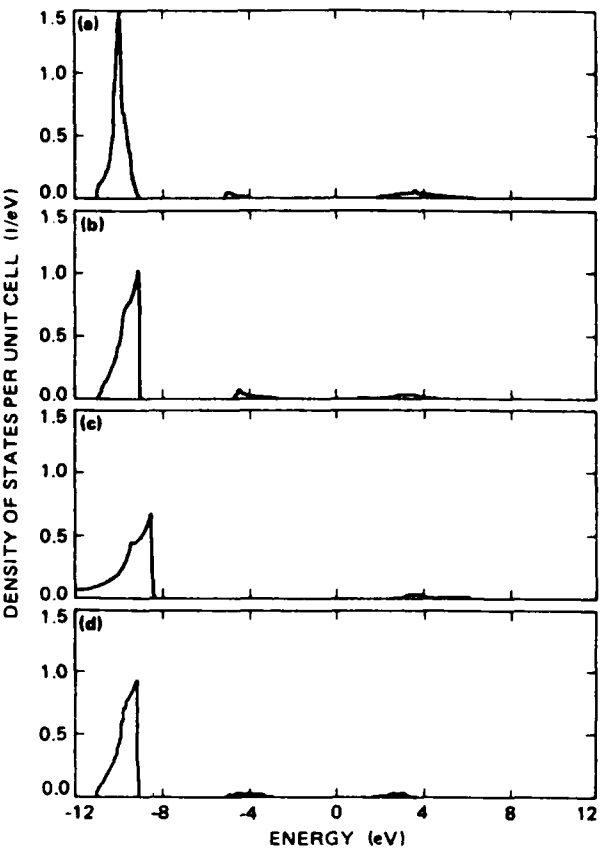


FIG. 5. The Te Γ_6 partial densities of states.

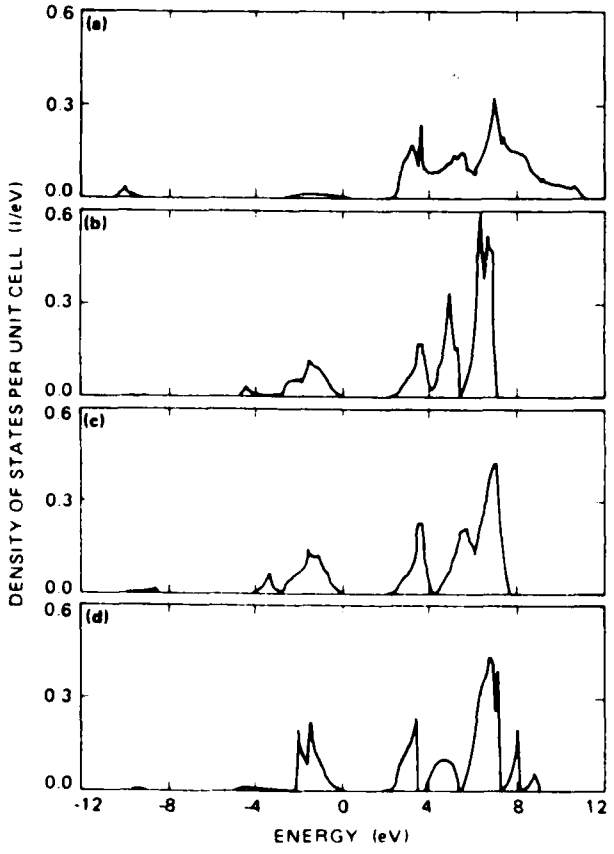


FIG. 4. The Cd Γ_7 partial densities of states.

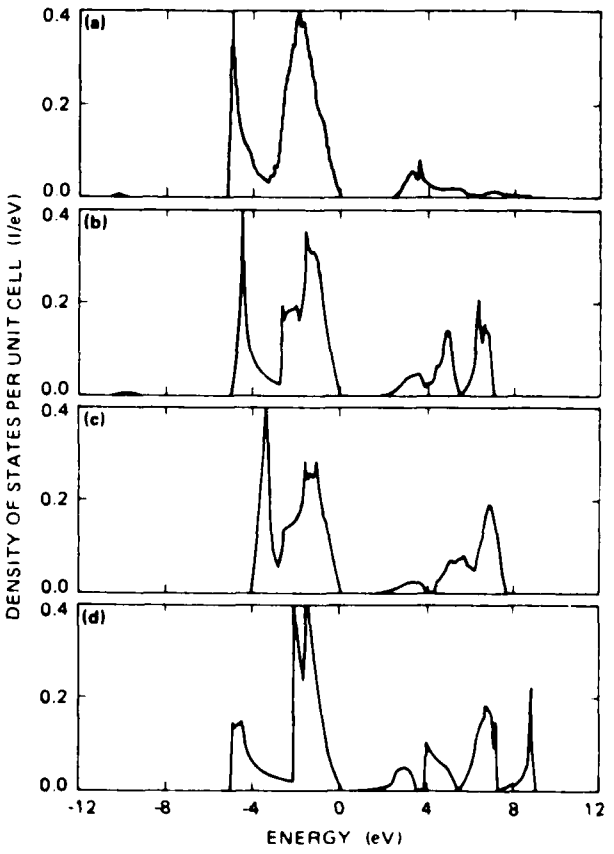


FIG. 6. The Te Γ_7 partial densities of states.

evolve into the conduction bands, while the anion p states make up most of the major valence bands just below the gap. Thus it is perhaps a surprise to see a prominent peak derived from the cation s states at the bottom of the major valence-band structure. However, this is a general feature for all sp^3-based compound semiconductors. These are the states responsible for the first observed breakdown of the virtual-crystal approximation for a semiconductor alloy: $Hg_{1-x}Cd_xTe$ (which is caused by the large s-energy shift between the Cd and Hg sites).[4,5,22]

A more detailed examination draws attention to some important differences among the four panels in Fig. 3: the valence-band peak in panel (c) is about 2 eV higher than the rest, and it is also high compared to experiment.[22] Our conduction-band PDOS in panel (a) is broader than the others. The ratio of the integrated PDOS in the conduction bands to that in the valence bands in our model is larger than those in other panels. Also our PDOS just below the valence-band edge is obviously smaller than that found in other models.

Figure 4 shows that the Cd p states are concentrated in the conduction-band states. This is particularly true in panel (a), where their contribution to the valence-band states shrinks almost to nothing. In other panels, there are still sizable ($\sim 20\%$) valence-band states. In contrast, all four panels in Fig. 5 show that the Te s states are confined to the deep valence-band states, as generally recognized. Finally, Fig. 6 shows that the Te p states dominate the upper valence-band states. Panel (a) has much less conduction-band content than the other three panels. As we will see, these differences can result in quantitatively or even qualitatively different predictions about the deep levels.

IV. IMPURITY-LEVEL DETERMINATION

A convenient way to study the impurity energy levels using Eq. (1) is to rewrite it as $v_\alpha = 1/g_\alpha(E)$ and plot E as a function of v. Once this E-v curve is deduced for each α, the deep levels E_α for a given impurity can be read off the curve by drawing a vertical line at the appropriate value of v_α for the impurity. We set the zero of energy at the top of the valence bands. Because the gap is 1.6 eV, we will focus on levels in the energy range from -0.5 to 2.0 eV.

Calculations have been performed for all neutral impurities listed in Table I. Because we do not believe that there exists a uniformly accepted table for v we have adopted a table that we used for structural studies.[23,24] Table I lists the term values, which we obtained from to-

TABLE I. s- and p-state correlated term values in units of $-$eV. The top entry is the s-state, the second the $p_{1/2}$-state, and the third the $p_{3/2}$-state energy. (All energies are *negative*.)

I	II	III	IV	V	VI	VII
Li	Be	B	C	N	O	F
5.390	9.320	14.003	19.814	26.081	28.551	36.229
	5.412	8.300	11.260	14.540	13.613	17.484
	5.412	8.300	11.260	14.540	13.610	17.420
Na	Zn	Al	Si	P	S	Cl
5.140	9.390	11.780	15.027	19.620	21.163	25.812
	4.237	5.980	8.150	10.610	10.449	13.136
	4.011	5.980	8.150	10.550	10.360	13.010
K	Cd	Ga	Ge	As	Se	Br
4.340	8.990	13.230	16.396	20.015	21.412	24.949
	4.313	6.000	7.880	10.146	10.188	12.353
	4.097	5.850	7.694	9.810	9.750	11.840
Rb	Hg	In	Sn	Sb	Te	I
4.180	10.430	12.032	14.525	17.560	19.120	21.631
	4.998	5.780	7.340	9.391	9.951	11.470
	4.031	5.453	6.879	8.640	9.010	10.450
Cs			Pb			
3.890			15.250			
			7.410			
			5.979			
Cu						
7.720						
Ag						
7.570						
3.647						
3.487						
Au						
9.220						
4.349						
3.688						

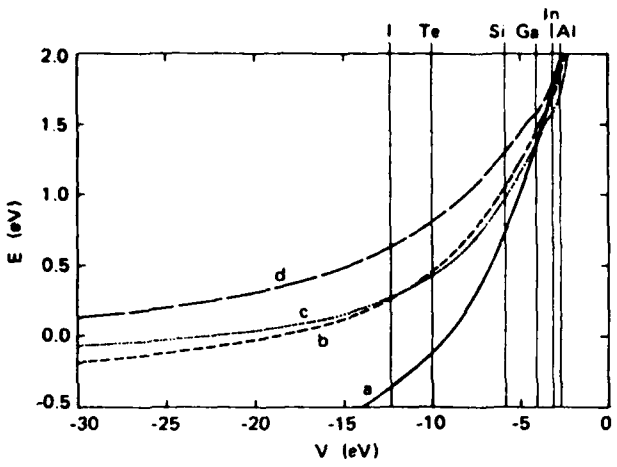


FIG. 7. The E-v curves for the Γ_6 states on a Cd site.

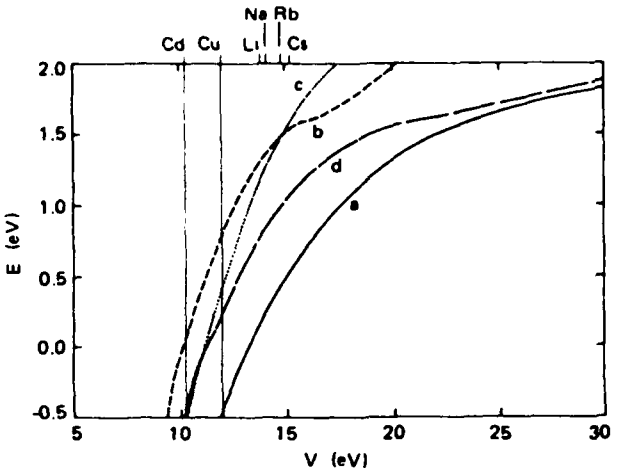


FIG. 9. The E-v curves for the Γ_6 states on a Te site.

tal energy differences between atomic configurations calculated using the norm-conserved pseudopotentials[25] and self-consistent charge-density-functional theory, with the first ionization energies adjusted to be the experimental values.[26] These term values are found to yield consistently better structural properties[23] in Harrison's theory[27,28] than those based on Mann's values[29] adopted by Harrison.[28] The impurity-potential parameters will then be taken as the difference of the term values between the impurity atom and Cd (or Te). To study the sensitivity of E_a to v_a, we shift v_a by ± 0.5 eV and compute the corresponding changes in the energy levels.

Figures 7–10 display the E-v curves for several α. Each figure has four curves, corresponding to the four panels of PDOS in each of Figs. 3–6. The functional behavior of these curves can be understood qualitatively using Eq. (2) and Figs. 3–6. If E lies in the gap, the contribution from conduction bands is negative, but positive from the valence bands. The closer the PDOS to the E in question, the larger will be its influence. Applying this argument to the Γ_6(Cd) representation, we see that the curves in Fig. 7 are negative in the gap region because the PDOS in Fig. 3 near the bottom of the conduction bands are much larger than those near the valence-band top. Thus, on the Cd site, only impurities with an s energy below the Cd s level (-8.99 eV) will produce a Γ_6 level in the gap. However, we note that in Fig. 7, $g_a(E)=0$ for models (b) and (c) just below the valence-band edge because of cancellation between the conduction- and valence-band contributions. At this E value, the E-v curve switches from $v=-\infty$ to $v=\infty$ (not shown); an ideal vacancy level (corresponding to $v_a=\infty$) is located at this E. A similar consideration, but with the conduction and valence bands interchanged, leads to an understanding of the curves in Fig. 10. Using the same principle, we can easily understand why all curves in Fig. 9 for the Γ_6(Te) representation are positive, but the reasons for the large displacements between these curves are not easy to deduce. In Fig. 8, the curve labeled a is distinctively different from other curves, because the PDOS in panel (a) in Fig. 4 is completely dominated by the conduction band; however, for the other panels the PDOS just below the valence-band edge are as large as those just above the conduction-band edge. This produces a very sharp negative E-v curve for (a), but split behavior for (b), (c), and (d).

These E-v curves provide a clear picture of how different host band structures may affect the deep levels. Numerical values for the impurity levels can be obtained from these figures by drawing vertical lines at the ap-

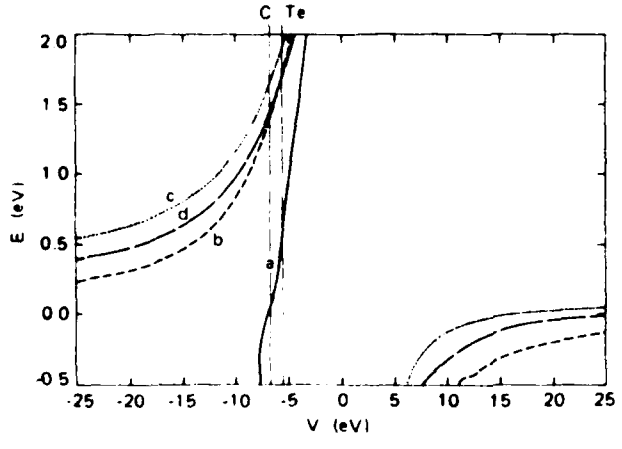


FIG. 8. The E-v curves for the Γ_6 states on a Cd site.

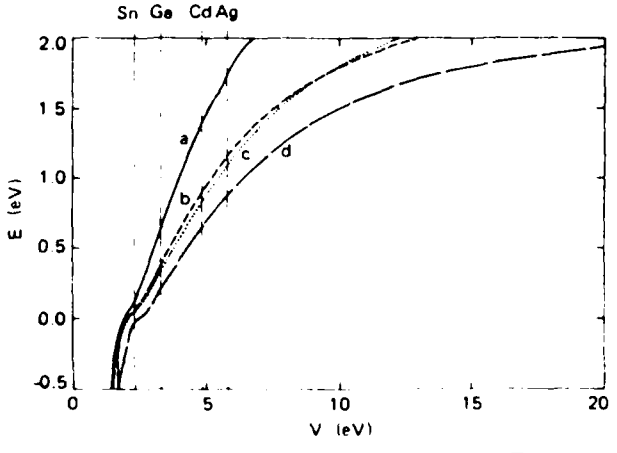


FIG. 10. The E-v curves for the Γ_6 states on a Te site.

propriate impurity potentials (i.e., differences between the term values listed in Table I), as has been shown for several representative impurities. To provide a more quantitative comparison, Table II lists some calculated impurity levels E_α and the corresponding changes ΔE_α due to the 1-eV change in v_α.

V. RESULTS AND CONCLUSION

To summarize we recall that band models (b) and (c) are the same second-neighbor ETB with two different sets of parameters, and model (d) is a first-neighbor ETB with one extra s orbital per atom. Our model [model (a)] has the form of ETB but is derived in a very different manner and includes all the long-range interactions. Therefore, we expect that the results from models (b) and (c) will be close, model (d) will have larger discrepancies from (b) and (c) than that between (b) and (c), and model (a) will differ even more. This is evident from Figs. 7–10 and Table II. We found the energies for the Γ_6(Cd), Γ_7(Te), and Γ_8(Te) states produced by models (b) and (c) agree within 0.1 eV. For the other states, i.e., Γ_6(Te), Γ_7(Cd), and Γ_8(Cd), the energies from (b) and (c) are qualitatively similar, but the difference can be as large as 0.4 eV. The largest discrepancy between models (d) and (b) [or (c)] is more than 0.5 eV, and that between (a) and other models is more than 1 eV. The largest difference comes from the p levels on a Cd site. For example, the filled p level of C on a Cd site in model (a) is a resonance state just below the valence-band edge but is a donor state in the other models. Similarly, model (a) puts the neutral Te antisite defect p levels at about $\frac{1}{3}$ and $\frac{2}{3}$ of the gap [$E(\Gamma_7)=0.48$ eV and $E(\Gamma_8)=0.95$ eV], while other models assign them as resonance states inside the conduction bands. We also note that the discrepancies between different models are not uniform, but vary with v_α. Consider Γ_6(Cd) for example. All four models yield the same ordering and about the same energies for the group-III impurities Al, In, and Ga. However, as v becomes more negative, the splitting between the curves increases, so the discrepancies become larger [~ 1 eV difference between models (a) and (d) for I impurity]. Similarly, for the Γ_7(Te) states, all four models put the Sn impurity energies close to the valence-band edge, but the agreement deteriorates as v_α increases.

Regarding the sensitivity of energy levels to impurity potentials, Table II shows that a 1-eV shift in v_α produces a change in E_α ranging from less than 0.1 to 0.65 eV. Very little is known about the size or trends in errors introduced in v_α from the use of atomic term values. However, we know that the discrepancy of v_α between two different tables of atomic term values can be larger than 2 eV. This discrepancy translates into an uncertainty of less than 0.1 to more than 1 eV in the impurity energy levels,

TABLE II. Defect energy levels E and changes ΔE due to a 1-eV change in the impurity-potential parameter. All energies are in units of eV. V_0 stands for ideal vacancy.

	Model (a)		Model (b)		Model (c)		Model (d)	
Defect	E	ΔE	E	ΔE	E	ΔE	E	ΔE
				Γ_6 on Cd site				
Ga	1.29	0.39	1.42	0.24	1.33	0.23	1.57	0.18
C	−0.21	0.09	0.38	0.09	0.36	0.13	0.74	0.08
Si	0.67	0.30	1.02	0.10	0.93	0.19	1.27	0.15
P	−0.19	0.11	0.39	0.09	0.38	0.08	0.75	0.08
O	< -0.5		−0.02	0.02	0.04	0.01	0.32	0.02
Te	−0.13	0.13	0.44	0.10	0.42	0.08	0.79	0.09
Cl	< -0.5		0.06	0.03	0.10	0.02	0.41	0.04
V_0	< 0.5		< -0.5		−0.30		−0.20	
				Γ_7 on Cd site				
C	−0.02	0.37	1.32	0.22	1.59	0.20	1.39	0.19
Si	1.57	0.65	> 2.0		> 2.0		> 2.0	
P	0.16	0.38	1.48	0.26	1.73	0.23	1.52	0.21
O	< 0.5		0.89	0.14	1.22	0.13	1.03	0.12
Te	0.48	0.55	1.60	0.29	1.88	0.23	1.66	0.24
Cl	< 0.5		0.96	0.17	1.29	0.14	1.09	0.14
V_0	< 0.5		0.00		0.21		0.06	
				Γ_6 on Te site				
Li	0.14	0.29	1.28	0.22	1.15	0.35	0.76	0.25
Cu	< 0.5		0.54	0.42	0.12	0.52	0.03	0.32
				Γ_7 on Te site				
Ag	1.89	0.32	1.26	0.22	1.21	0.23	0.99	0.20
Cd	1.66	0.34	1.11	0.26	1.05	0.26	0.85	0.22
Ga	0.98	0.49	0.61	0.33	0.55	0.32	0.40	0.30
Si	0.07	0.40	0.11	0.36	−0.13	0.38	0.38	0.72
Sn	0.28	0.47	0.15	0.31	0.13	0.24	0.02	0.28

which is comparable to that due to different host band structures.

Putting this large uncertainty in the deep levels against a band gap of 1.6 eV, we are left with great doubts about the predictability of this oversimplified theory. Unfortunately, the experimental means available for identifying microdefects in semiconductors are still very limited, and the *ab initio* band theory is still not capable of accurately predicting the energy levels. Thus, there is a great temptation to use simple theories like the one carried out here to help with the identifications. To illustrate this point, consider the following examples: Table II shows that Li on a Te site has an s level of 0.14 eV in model (a), so one may be tempted to relate it to the acceptor state identified experimentally.[14] However, this is not the hydrogenic acceptor state on a Cd site, as one might anticipate. One might also want to assign the $\frac{1}{3}$ and $\frac{2}{3}$ gap states for the Te antisite p levels on the Cd site found from model (a) as those seen in experiments.[15,16] Because of the large uncertainty in the calculation, these results should be regarded as suspicious surprises rather than theoretical confirmations.

The results presented here should not discourage continued research on the ETB approach, but improvement is clearly needed. Work ranging from universal[23,27,30] to specific[24,31,32] structural studies to our band calculations and alloy studies[1-7] indicates that the ETB type of theory is practical for both bonding properties and electronic structures. The reason that ETB works well for some properties, e.g., photoemission spectra and bonding properties, but not for impurity levels, is that the former depend only on the gross total density of states, while the latter have been shown to be sensitive to the details of the partial densities of states.

To establish the credibility of ETB in defect studies, one needs to look at the problem more seriously. The most difficult and yet important task is to develop a better way for determining the Hamiltonian matrix elements. Haas *et al.*[8] and Harrison[27,28] have suggested using the atomic term values as the diagonal matrix elements. Our work[1-4] has suggested using a universal long-range interaction to improve the accuracy of the conduction bands. Several studies[1,27,28,33] have also pointed out scaling rules of the matrix elements. A combination of these ideas may lead to an acceptable model. Secondly, both the bonding and deep-level states of impurities should be studied at the same time in order to provide correlated information for defect identification. Finally, more realistic models should be examined. Besides the substitutional site-diagonal defects, one should consider the possibility of interstitial, paired, and even more complex defects. One also needs to deal with long-range impurity potentials, possible charge shifts, and lattice distortions. Progress in all these areas can be expected if the calculation is constantly correlated with experiments and available *ab initio* theory.

ACKNOWLEDGMENTS

This work was supported by AFOSR Contract No. F49620-81-K0012 and Grant No. AFOSR-84-0282. A.-B. Chen would like to thank Professor W. E. Spicer of Stanford University for his hospitality.

[1] A.-B. Chen and A. Sher, Phys. Rev. B 22, 3886 (1980).
[2] A.-B. Chen and A. Sher, Phys. Rev. B 23, 5645 (1981).
[3] A.-B. Chen and A. Sher, Phys. Rev. B 23, 5360 (1981).
[4] A.-B. Chen and A. Sher, J. Vac. Sci. Technol. 121, 138 (1982).
[5] W. E. Spicer, J. A. Silberman, J. Morgan, J. Lindau, J. A. Wilson, A.-B. Chen, and A. Sher, Phys. Rev. Lett. 149, 948 (1982).
[6] A.-B. Chen and A. Sher, Phys. Rev. B 26, 6603 (1982).
[7] A.-B. Chen, S. Phokaichaipatana, and A. Sher, Phys. Rev. B 23, 5360 (1981).
[8] K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B 27, 5360 (1983).
[9] C. A. Swarts, M. S. Daw, and T. C. McGill, J. Vac. Sci. Technol. 121, 199 (1982).
[10] A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 25, 6367 (1982).
[11] J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. B 21, 3545 (1980).
[12] G. A. Baraff, E. O. Kane, and M. Schluter, Phys. Rev. B 21, 5662 (1980).
[13] G. B. Bachelet, M. Schluter, and G. A. Baraff, Phys. Rev. B 27, 2545 (1983).
[14] K. Zanio, in *Semiconductors and Semimetals*, edited by R. Willardson and A. C. Beer (Academic, New York, 1978), Vol. 13.
[15] C. E. Jones, V. Nair, J. Lundquist, and D. L. Polla, J. Vac. Sci. Technol. 21, 187 (1982).
[16] R. T. Collins and T. C. McGill, J. Vac. Sci. Technol. A1, 1633 (1983).
[17] K. K. Kanazawa and F. C. Brown, Phys. Rev. 135, A1757 (1964).
[18] A.-B. Chen, Phys. Rev. B 16, 3291 (1977).
[19] D. J. Chadi, J. P. Walter, and M. L. Cohen, Phys. Rev. B 5, 3058 (1972).
[20] E. O. Kane, Phys. Rev. B 13, 3478 (1976).
[21] D. J. Chadi, Phys. Rev. B 16, 3572 (1977).
[22] J. A. Silberman, P. Morgan, W. E. Spicer, and J. A. Wilson, J. Vac. Sci. Technol. 21, 142 (1982).
[23] A.-B. Chen and A. Sher, Microscience 3, 1 (1984).
[24] A. Sher, A.-B. Chen, and W. E. Spicer, in *Thirteenth International Conference on Defects in Semiconductors*, edited by L. C. Kimerling and J. M. Parsey, Jr. (The Metallurgical Society of AIME, 1985), p. 335.
[25] G. B. Bachelet, D. Hamman, and M. Schulter, Phys. Rev. B 26, 4199 (1982).
[26] C. Kittel, *Introduction to Solid State Physics, 5th edition* (Wiley, New York, 1976), p. 75, Table 2.
[27] W. A. Harrison, *Electronic Structure and Properties of Solids*

(Freeman, San Francisco, 1980).
[28]W. A. Harrison, Microscience 3, 35 (1983).
[29]J. B. Mann, *Atomic Structure Calculations, I: Hartree-Fock Energy Results for Elements Hydrogen to Lawrencium* (Clearinghouse for Technical Information, Springfield, Virginia, 1967).
[30]D. G. Pettifor and R. Podloucky, Phys. Rev. Lett. 53, 826 (1984).
[31]D. C. Allan and E. J. Mele, Phys. Rev. Lett. 53, 826 (1984).
[32]D. J. Chadi, Phys. Rev. Lett. 52, 1911 (1984).
[33]O. K. Andersen, W. Klase, and M. Nohl, Phys. Rev. B 17, 1209 (1978).

Generalized Brooks' formula and the electron mobility in $Si_x Ge_{1-x}$ alloys

Srinivasan Krishnamurthy and A. Sher
SRI International, Menlo Park, California 94025

An-Ban Chen
Auburn University, Auburn, Alabama 36830

(Received 19 March 1985; accepted for publication 7 May 1985)

A formula for alloy-scattering-limited electron mobility in semiconductors is obtained for indirect gap systems with multiple band minima. All the input parameters needed are defined explicitly. The drift mobility of $Si_x Ge_{1-x}$ which has a dip at $x \sim 0.13$ and a broader minimum at $x \sim 0.5$ is calculated by adding alloy scattering to other scattering mechanisms and correlates well with the measured Hall mobility.

The electron and hole mobilities in semiconductors are determined by the band structure and various scattering mechanisms, predominately impurity and phonon scattering. For alloys, the mobility is also affected by disorder arising from aperiodic atomic potentials and atomic positions. Many years ago, Nordheim[1] and Brooks[2] obtained an expression for alloy-scattering-limited electron mobilities in metals and semiconductors, respectively. Brooks' well-known formula reads

$$\mu_A = \frac{\sqrt{2\pi}e\hbar^4 N_0}{3x(1-x)m^{*5/2}} \frac{1}{(\Delta E)^2\sqrt{KT}}, \tag{1}$$

where N_0 is the number of atoms per unit volume, m^* is a band-edge effective mass, x is the fractional concentration of one of the species, and ΔE is an energy parameter characterizing the alloy potential fluctuations. Although this formula has been widely and, to some extent, successfully used for direct gap materials,[3-5] the identification of the alloy disorder parameter ΔE remains uncertain. Various suggestions have previously been made for ΔE, e.g., and band-edge discontinuity[6] or band-gap differences.[7] Any of these simple choices is bound to fail when one applies Eq. (1) to more complicated indirect gap systems such as $Si_x Ge_{1-x}$ alloys, where one encounters conduction-band minima transferring between the X and L points of the Brillouin zone. For example, if ΔE is taken to be the difference in corresponding band edges, then one finds that $\Delta E \sim 0.1$ eV for the X (Δ) valley and ~ 1.2 eV for the L valley. The values that fit the experiment are about half this value for L and ~ 0.5 eV for X.[6] The purpose of this letter is to resolve the identity of ΔE for indirect gap materials.

Moreover, there is a problem with the m^* that enters Eq. (1). For direct gap alloys, the band-edge effective mass at Γ naturally enters Eq. (1). For the indirect gap alloys, the effective mass is anisotropic and hence an appropriate mass must be chosen. Previous authors[6,7] have chosen m^* to be the effective conductivity mass m_c^*. We shall show that different masses enter for different cases.

The first unambiguous assignment for ΔE in a direct gap alloy was given by Hass *et al.*[4] To estimate the limiting electron mobility in $Hg_x Cd_{1-x} Te$ based on a tight-binding (TB) band description, they defined ΔE to be $f_s \Delta E_s$, where f_s is the s fraction in the density of states and ΔE_s is the difference between the s atomic term values of the Hg and Cd atoms. By extending this approach to alloys with indirect gaps and multiple bands, we show that all the uncertainties identified above are resolved. Our generalized Brooks' formula will then be applied to $Si_x Ge_{1-x}$ systems to explain their observed mobility.[7,8]

Because Brooks' formula has never been derived explicitly in the literature, we rederive it first and then generalize it. Consider the case of a single band with an isotropic effective mass. The dc electronic conductivity based on the linear response theory[9] is given by

$$\sigma = \int \sigma(\epsilon)\left(\frac{-df}{d\epsilon}\right)d\epsilon, \tag{2}$$

where the energy-dependent $\sigma(\epsilon)$ in the weak alloy scattering limit is

$$\sigma(\epsilon) = (e^2/3)v^2(\epsilon)D(\epsilon)\tau(\epsilon). \tag{3}$$

$D(\epsilon)$ is the density of states (DOS) per unit volume for both spins, so $D(\epsilon) = 2N_A\rho(\epsilon)$, with $N_A = N_0/2$ being the number of unit cells per volume (for the diamond structure, half the number of atoms N_0 per unit volume) and $\rho(\epsilon)$ being the DOS per unit cell per spin. The mean square velocity $v^2(\epsilon)$ for carriers with energy ϵ is given by

$$v^2(\epsilon) = \sum_{\mathbf{k}} |v(\mathbf{k})|^2 \frac{\delta[\epsilon - \epsilon(\mathbf{k})]}{\rho(\epsilon)}. \tag{4}$$

The scattering lifetime for carriers with energy ϵ, $\tau(\epsilon)$, is related to the alloy broadening $\Delta(\epsilon)$ by $\tau(\epsilon) = \hbar/2\Delta(\epsilon)$, where the energy $\Delta(\epsilon)$ is the imaginary part of the self-energy in the averaged alloy Green's function. For weak scattering $\Delta(\epsilon)$ is

$$\Delta(\epsilon) = \pi x(1-x)(\Delta E)^2\rho(\epsilon), \tag{5}$$

where in a tight binding (TB) description ΔE is the difference in the term values of the constituents. Then the mobility is $\mu_A = \sigma/ne$ with the electron density given by

$$n = 2\int f(\epsilon)\rho(\epsilon)d\epsilon. \tag{6}$$

For a nondegenerate semiconductor, $f(\epsilon)$ is the Boltzmann distribution and $f(\epsilon) \propto e^{-(\epsilon-\epsilon_f)/kT}$. Furthermore, for a parabolic band $\epsilon(\mathbf{k}) = \hbar^2k^2/2m^*$, $\rho(\epsilon) = (2m^*)^{3/2}\epsilon^{1/2}/4\pi^2\hbar^3$, then all the above equations can be combined to arrive at Eq. (1).

For a real semiconductor alloy in a TB description, the alloy scattering can be characterized by two parameters ΔE_s and ΔE_p, the differences in s and p term values between two substitutional atoms. Then an effective broadening is given

 0003-6951/85/140160-03$01.00

by

$$\Delta(\epsilon) = (\Delta_s \rho_s + \Delta_p \rho_p)/\rho, \tag{7}$$

where ρ_s, ρ_p are partial density of states (PDOS) and Δ_s and Δ_p are similar to Eq. (5), with ρ replaced by ρ_s and ρ_p, respectively. For $Hg_{1-x}Cd_xTe$, the s disorder is predominent[5,10] and one can neglect ΔE_p. Defining $\rho_s = f_s \rho$ (and $\rho_p = f_p \rho$), one arrives at

$$\Delta \approx \pi x(1-x)(f_s \Delta E_s)^2 \rho.$$

Thus as was pointed out by Hass *et al.*,[5] $f_s \Delta E_s$ plays the role of ΔE in this special case where ΔE_p can be neglected.

For an alloy with a single indirect gap minimum, one has to consider both s and p contributions to the alloy broadening and the masses that enter ρ and v^2. Again, Eqs. (2)–(6) can be combined to yield

$$\mu_A = \frac{(e\hbar^4 N_0 \sqrt{2\pi})}{[3x(1-x)m_c^* m_t^* (m_l^*)^{1/2}(kT)^{1/2} N_v (f_s^2 \Delta E_s^2 + f_p^2 \Delta E_p^2)]}, \tag{8}$$

where m_l^* and m_t^* are respectively the longitudinal and the transverse mass at the band edge, and N_v is the number of equivalent minima, e.g., 6 for Si. The conductivity mass m_c^* comes from averaging v^2 in Eq. (2) and is given by $3(2/m_t^* + 1/m_l^*)^{-1}$. Equation (8) clearly identifies the masses and the energy parameter that enter Brooks' formula.

Next we consider a still more complicated case where the contribution to the mobility comes from more than one band. For example, in Si_xGe_{1-x} the X and L minima cross near $x = 0.15$.[11] There are now two contributions to the net conductivity, so $\sigma = \Sigma \sigma_i$, where i is X or L. The quantities $v_i^2(\epsilon)$, $D_i(\epsilon)$, and $N_i(\epsilon)$ now take different values for different bands. The structure of $\tau_i(\epsilon)$ requires more careful consideration. The complication comes from the fact that the effective broadening Δ is still given by Eq. (7), but ρ_s, ρ_p, and ρ contain contributions from both the bands. The proper expressions are $\rho = \Sigma_i \rho_i N_v^i$ and $\rho_\alpha = \Sigma_i f_{\alpha i} N_v^i \rho$, where $i = X$ or L, $\alpha = s$ or p, and $N_v^X = 6$, $N_v^L = 4$. The equation for Δ is

$$\Delta(\epsilon) = \pi x(1-x) \sum_\alpha \left(\sum_i f_{\alpha i} N_v^i \rho_i(\epsilon) \Delta E_\alpha \right)^2 \Big/ \left(\sum_i \rho_i(\epsilon) N_v^i \right). \tag{9}$$

The mobility associated with the ith band is defined as $\mu_i = \sigma_i/(n_i e)$, then

$$\mu_i^A = \frac{\pi e \hbar^4 N_0}{3x(1-x)} \frac{1}{[m_c^* m_t^* (2m_l^*)^{1/2}]_i} I_i, \tag{10}$$

$$I_i = \int_0^\infty \frac{\epsilon \rho_i(\epsilon) \left(\sum_j N_v^j \rho_j(\epsilon) \right) e^{-\epsilon/kT}}{\sum_\alpha \left(\sum_j f_{\alpha j} N_v^j \rho_j(\epsilon) \Delta E_\alpha \right)^2} d\epsilon \Big/ \int \epsilon^{1/2} e^{-\epsilon/kT} d\epsilon. \tag{11}$$

Thus, the generalized formula no longer has the explicit x and T dependences of the original Brooks' form. However, all the quantities needed—the masses, the scattering parameters ΔE_α, the band gaps, and the fractions $f_{\alpha i}$—can be evaluated theoretically without resorting to experimentally fitted parameters. To demonstrate, we shall apply Eq. (10) to Si_xGe_{1-x}. The band quantities are obtained from our CPA calculation.[11] We found that the effective masses vary weakly with the concentration, so m_l^* and m_t^* are assumed to be constant and assigned the values 0.97 and 0.19 for the X minima and 1.64 and 0.082 for the L minima, respectively. The calculated energy gaps for the $X(\Delta)$ follows the functional form $E_g^{(X)} = a + bx + cx^2$ and for L is given by $E_g^{(L)} = A + Bx + Cx^2$. All the parameters of our calculations are listed in Table I.

TABLE I. Calculation parameters.

Parameter	Si_xGe_{1-x} systems
$m_l^*(X)$	$0.97m_0$
$m_t^*(X)$	$0.19m_0$
$m_l^*(L)$	$1.64m_0$
$m_t^*(L)$	$0.082m_0$
$E_g^X(x)$	$0.8941 + 0.0421x + 0.1691x^2$
$E_g^L(x)$	$0.7596 + 1.0860x + 0.3306x^2$
$f_{sX}(x)$	$0.333 + 0.05x$ $(0<x<0.3)$
	$0.339 + 0.03x$ $(0.3<x<1.0)$
$f_{sL}(x)$	$0.632 + 0.13x$

To correlate the calculation with the measured mobilities, we need to have an estimate of scattering rates $1/\tau_0$ due to impurities and phonons. A crude approximation is to assume $1/\tau_0$ for a given valley to be the same as the appropriate constituent's values and add to it the alloy scattering rate $1/\tau_A$. Then the average mobility and the mobility from the ith minimum in the alloy are

$$\mu = \sum_i n_i \mu_i \Big/ \sum_i n_i,$$

$$(\mu_i)^{-1} = (\mu_i^0)^{-1} + (\mu_i^A)^{-1}, \tag{12}$$

μ_i^A is given by Eq. (10) and μ_i^0 are the measured drift mobilities for Si or Ge.[12] The drift mobility, calculated from Eq. (12), is plotted as a function of alloy concentration x in Fig. 1.

For $x<0.05$ and $x>0.20$, the energy difference between the X and L edges is large enough so there is a negligible contribution to the mobility from the higher minima. In the Si_xGe_{1-x} system, the s scattering is predominent. Because

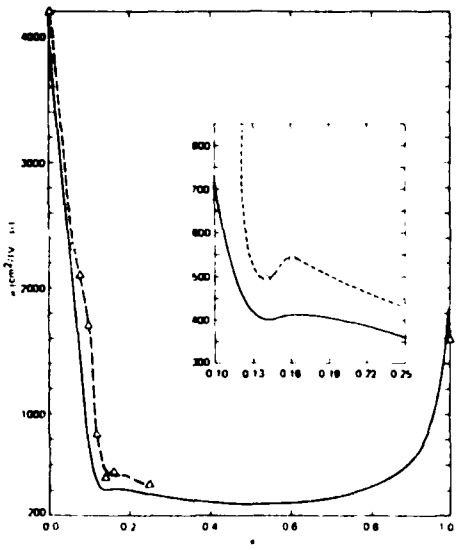


FIG. 1. Calculated (solid line) electron drift mobility and the experimental Hall mobility (dashed line) from Refs. 6 and 8 are plotted as a function of alloy concentration.

the L edges have a larger s content, alloy scattering reduces the average mobility substantially for small x. Even though the s content is almost the same for all $x > 0.20$ at the X edge, the mobility still decreases to $x \simeq 0.5$ as shown in Fig. 1 because of the $x(1-x)$ term in Eq. (10).

An interesting feature is obtained for the compositions $0.13 < x < 0.18$. The average mobility attains a local minimum near $x = 0.14$ and a smaller maximum at $x = 0.17$. This feature occurs because of the X to L crossover.[11] For $x < 0.14$, the major contribution to μ comes from L minima. Near $x = 0.14$, the density of states increases because the X and L minima merge. So the alloy scattering increases there and the average mobility decreases. For $x > 0.14$, the X bands have the lower minima. As the s content is small at the X minima, the reduced alloy scattering increases the average mobility. For larger values of x, the $x(1-x)$ term takes over and the mobility varies as shown. The values of measured Hall mobility in $Si_x Ge_{1-x}$ systems are also plotted in Fig. 1. The interesting feature near $x = 0.14$ is clearly seen. Since the experimental drift mobility μ_D for $Si_x Ge_{1-x}$ is not available and the generalization of Eq. (10) to Hall mobility μ_H is less clear, we present the calculated μ_D and experimental μ_H (Ref. 7,8) here. While we do not expect quantitative agreement, because μ_H/μ_D can range from 1 to 2,[13,14] we do expect them to display the same qualitative x dependence. It is rewarding to note the similarity in the trend in Fig. 1. Previous authors explained the dip in the mobility curve by including intervalley scattering with an arbitrary adjustable coupling constant.[6] Our calculations automatically include that portion of intervalley scattering that results from alloy disorder with a coupling constant set by the atomic properties of the constituents. However, the additon of intervalley scattering mediated by phonons and impurities is expected to increase the dip near crossover.

In summary, an expression for alloy-scattering-limited charge carrier mobilities is derived for indirect gap alloys with multiple bands. This expression reduces to Brooks' formula for direct gap alloys. The quantities m^* and ΔE can be calculated exactly. Alloy scattering accounts for the observed mobility features in the $Si_x Ge_{1-x}$ alloy, including the anomaly near the L to X (Δ) crossover.

A.-B. would like to thank Professor W. E. Spicer for his hospitality at Stanford University. This work was supported in part by DARPA contract MDA 903-83-C-0108 and grant AFOSR-84-0282.

[1] L. Nordheim, Ann. Phys. **9**, 607 (1931); **9**, 641 (1931).
[2] H. Brooks (unpublished). A discussion of this formula can be found, for example, in L. Makowski and M. Glicksman, J. Phys. Chem. Solids **34**, 487 (1973).
[3] A. Chandra and L. F. Eastman, J. Appl. Phys. **51**, 2669 (1980).
[4] D. Chattopadhyay and B. R. Nag, Phys. Rev. B **12**, 5676 (1975).
[5] K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B **27**, 1088 (1983).
[6] J. W. Harrison and J. R. Hauser, J. Appl. Phys. **47**, 292 (1976).
[7] M. Glicksman, Phys. Rev. **111**, 125 (1958).
[8] M. Glicksman, Phys. Rev. **100**, 1146 (1955).
[9] A.-B. Chen, G. Weisz, and A. Sher, Phys. Rev. B **5**, 2897 (1972). See Eq. (134).
[10] D. S. Montgomery, J. Phys. C **16**, 2923 (1983).
[11] S. Krishnamurthy, A.-B. Chen, and A. Sher (unpublished).
[12] S. M. Sze, *Physics of Semiconductors*, 2nd ed. (Wiley–Interscience, New York, 1981).
[13] V. A. Johnson and K. L. Horowitz, Phys. Rev. **79**, 176 (1950); **79**, 409 (1950).
[14] H. Jones, Phys. Rev. **81**, 149 (1951).

Binding Energy and Spectral Width of Si $2p$ Core Excitons in Si_xGe_{1-x} Alloys

S. Krishnamurthy and A. Sher
Physical Electronics Laboratory, SRI International, Menlo Park, California 94025

and

A.-B. Chen[a]
Department of Physics, Auburn University, Auburn, Alabama 36849
(Received 24 January 1985)

A calculation is presented to explain the anomalous experimental behavior of the Si $2p$ core-exciton binding energy and linewidth in Si_xGe_{1-x} alloys. The observed minimum in the linewidth near $x \simeq 0.15$ can be explained as the result of a competition between intrinsic broadening due to screening and extrinsic alloy broadening. For pure Si, the binding energy is estimated to be 0.15 ± 0.05 eV and the width is shown to be smaller than that observed at $x \simeq 0.15$.

PACS numbers: 71.55.Fr, 71.35.+z, 78.70.Dm

Until 1984, the Si $2p$ core exciton was believed to have an anomalously large binding energy.[1-10] Later, Newman and Dow[11] proposed a radically different picture in which the Si $2p$ core exciton is in fact a resonance with a negative binding energy. They further predicted that the exciton binding energy remains negative throughout most of the Si_xGe_{1-x} alloy composition range, except near $x \simeq 0.20$ where it becomes positive. In a recent experiment,[12] Bunker *et al.* found an anomalous sharpening of the exciton spectra near $x = 0.15$; the data were interpreted to support the Newman-Dow point of view. Yet the most recent experiment[13] still suggests a positive value for the binding energy E_b^0 in silicon.

In this Letter, we present a calculation that offers a plausible resolution to the above problem. In our theory, the calculated Si $2p$ core-exciton binding energy $E_b(x)$ and the linewidth $\Delta(x)$ in Si_xGe_{1-x} alloys are sensitively dependent on the parameter E_b^0. A comparison of the calculated $\Delta(x)$ with the experiment[12] suggests a positive value 0.15 ± 0.05 for E_b^0. The anomalous experimental spectrum[12] near $x \simeq 0.15$ is explained as a result of a competition between an intrinsic broadening Δ_I due to screening and an extrinsic alloy broadening Δ_A. In the present theory, there is no need to suppose that the exciton suddenly changes its character from an extended effective-mass-like state to a deep localized state.

We need to calculate E_b and $\Delta = \Delta_I + \Delta_A$ as a function of alloy concentration x. The calculations are based on a quantitative coherent-potential-approximation (CPA) band structure. Details of the CPA calculations will be presented elsewhere. Below, we briefly discuss a Green's-function method for calculating E_b and Δ_A.

The one-particle effective Green's function in CPA takes the form

$$G(E) = [E - \bar{H} - \Sigma(E)]^{-1}, \tag{1}$$

where $\bar{H}$ is the virtual-crystal approximation Hamiltonian and $\Sigma(E)$ is the self-energy. The site-diagonal Green's function is denoted as

$$F_\alpha(E) = \langle \phi_\alpha | G(E) | \phi_\alpha \rangle, \tag{2}$$

where ϕ_α is a localized orbital of specified symmetry. Here we only need to consider $\alpha = s$ for A_1 symmetry. The corresponding function in pure Si is denoted as $F_s^0(E)$. Following the theoretical treatment of deep substitutional-impurity levels,[8] the core-exciton level for pure Si is determined by

$$F_s^0(E) = (V - E_s^{Si})^{-1}, \tag{3}$$

where E_s^{Si} is the site potential seen by an s electron in bulk silicon, and V is a central-potential parameter. For a chosen value of V_b, Eq. (3) can be solved for E, and vice versa. Then $E_b^0 = E_c^0 - E$, where E_c^0 is the conduction-band edge in pure silicon. Because of the uncertainties in the value of experimental E_b^0 and theoretical V, we treat E_b^0 (or V) as a parameter. The binding energy E_b in a Si_xGe_{1-x} alloy can be calculated by solving

$$F_s(E) = [V - \bar{E}_s - \Sigma_s(E)]^{-1}, \tag{4}$$

where

$$\bar{E}_s = xE_s^{Si} + (1-x)E_s^{Ge}. \tag{5}$$

Then E_b is given by

$$E_b = E_c - E. \tag{6}$$

The calculated values of the conduction-band edge and the exciton level measured relative to the top of the valence band are plotted in Fig. 1. The band gap increases with x with a slope discontinuity at $\simeq 0.15$. The dashed lines a, b, and c represent exciton levels obtained with $E_b^0 = 0.1$, 0.15, and 0.30 eV, respectively. The binding energy E_b is also an increasing function of x, with a slope discontinuity near $x \simeq 0.15$. The CPA introduces a slight bowing in E_g and E_b.

Strinati[10] has calculated the variation of Δ_I with E_b

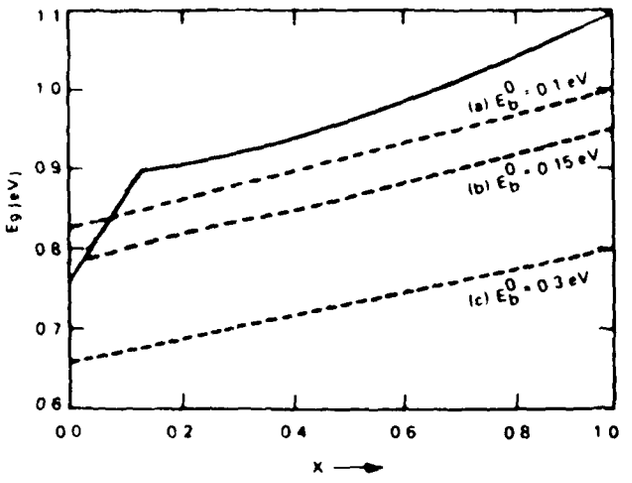


FIG. 1. Variation of the band gap (solid line) and the Si $2p$ core-exciton level with x in Si_xGe_{1-x} alloys. The energy is measured from the top of the valence band. The dashed curves a, b, and c represent exciton levels calculated with $E_b^0 = 0.1$, 0.15, and 0.3 eV, respectively.

by replacing the short-range Coulomb potential with a spherical square well of variable depth and a screened Coulomb tail. Strinati's results can be used to estimate Δ_I corresponding to the calculated E_b. Δ_I decreases rapidly with E_b, then saturates for larger E_b.

The contribution to the natural linewidth from the alloy broadening is calculated by a consideration of the electron part of the exciton wave function, ψ_s. The ψ_s is expanded in a linear combination of the s part of the conduction-band wave functions ϕ_n^s:

$$\psi_s(\mathbf{k}) = \sum_{n,s} C_{ns}\phi_n^s(\mathbf{k}). \tag{7}$$

We found that alloy scattering is only moderate and s scattering is dominant; thus, the alloy broadening $\Delta_4(E)$ is well approximated by

$$\Delta_4(E) = x(1-x)\delta_s^2 \operatorname{Im} F_s(E), \tag{8}$$

where δ_s is the difference between E_s^{Si} and E_s^{Ge}. Hence, the alloy-broadening contribution to Δ is related to the alloy broadening of the band states, $\Delta_n(\mathbf{k}, E)$:

$$\begin{aligned}\Delta_4 &= \frac{1}{N}\sum_{\mathbf{k}} \langle \psi_s(\mathbf{k}) | \Delta_4(E) | \psi_s(\mathbf{k}) \rangle \\ &= \frac{1}{N}\sum_{\mathbf{k}}\sum_{n} C_{n,s}^2 \Delta_n(\mathbf{k}, E) \\ &= \int \rho_s(E)\Delta_4(E)\,dE \\ &= x(1-x)\delta_s^2 \pi \int \rho_s^2(E)\,dE \end{aligned} \tag{9}$$

The integral in Eq. (9) is evaluated numerically.

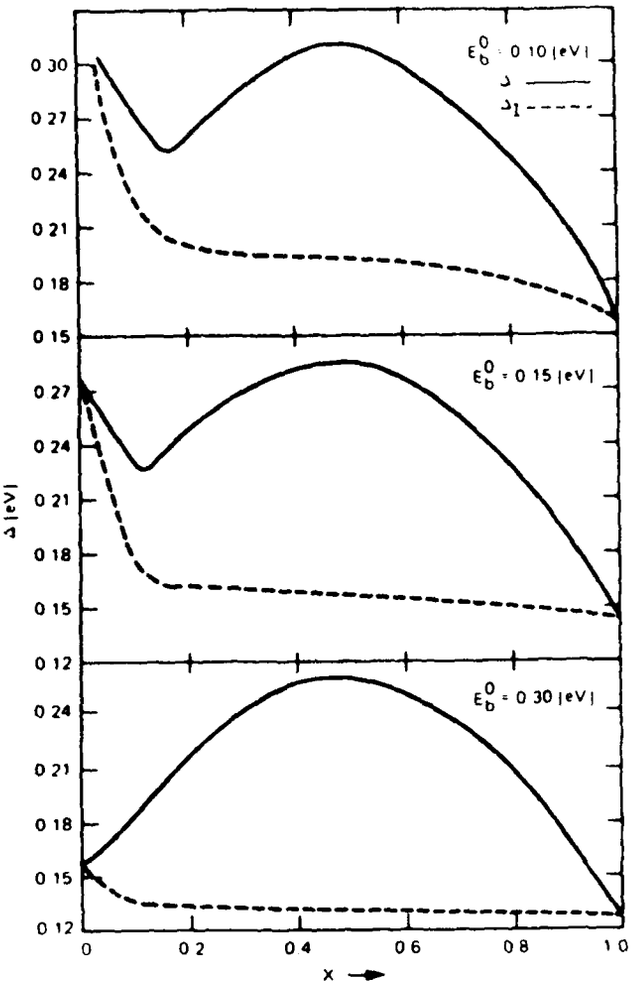


FIG. 2. Variation of Δ (solid lines) and Δ_I (dashed lines) with x for three E_b^0 values.

The calculated Δ, which is the sum of Δ_I and Δ_4, is plotted against x in Fig. 2 for three values of E_b^0. In all three panels, the dashed curve represents Δ_I and the solid line represents Δ. It is seen from Fig. 1 that the exciton level follows the X edge of the conduction band. Hence the binding energy E_b, relative to the conduction band edge, remains almost constant (for a given E_b^0) until the minimum switches from the X edge to the L edge. Because of the change in the slope of E_g, E_b decreases rapidly when L becomes the minimum. Correspondingly, Δ_I varies slowly until the X to L crossover and then increases rapidly. This feature is clearly seen in Fig. 2.

For $E_b^0 = 0.15$, the Δ_I and Δ_4 are comparable near $x = 0.50$, and Δ_I dominates for all small x and large x. These two competing mechanisms give a relative minimum near $x \approx 0.15$, a broader maximum near $x \approx 0.50$, and a smaller minimum for pure silicon. As E_b^0 is decreased, the relative minimum is shifted to larger x, e.g., the minimum shifts to $x = 0.20$ for $E_b^0 = 0.10$ eV. For $E_b^0 = 0.15$ eV, the position of the

relative minimum is in agreement with the experiment.[12] (By measuring the relative width at $x \approx 0.15$ to that $x = 1$, one can make a better estimate of E_b^0.) To correlate the theory with experiment, the calculated $1/\Delta^2$ is compared with the measured[12] $(\Delta\mu)^{-1}(d\mu/dE)_{max}$ in Fig. 3, where $\Delta\mu$ is the edge step and $(d\mu/dE)_{max}$ are the maximum values of the derivative of absorption spectra with respect to photon energy. Because the experimental values are given in arbitrary units, the values are normalized to agree at $x = 0.5$. The observed anomalous behavior near $x = 0.15$ and the qualitative x dependence in that region is clearly replicated by the theory. However, the calculation predicts a larger maximum at $x = 1$. It would be interesting to have experiments that cover the entire range of x to further test this prediction.

For larger values of E_b^0, the calculated E_b is also large and hence Δ_1 decreases slowly with x. Because the broadening is determined mainly by Δ_1, the linewidth is expected to be small for $x = 0$ and $x = 1$ or x; this occurs for $E_b^0 = 0.30$ eV. For negative values of E_b^0, E_b remains negative for all values of x. Accordingly, the linewidth is broad for all x, and there would be no such anomaly as in Fig. 3.

The calculations presented in this Letter are slightly different from alchemy approximations.[8] We treat the central-cell potential Δ as a parameter and narrow its range from other considerations. We examine values of $\Delta - E_s^{Si}$ of -8.49, -7.09, and -6.56 eV, corresponding to E_b^0 values of 0.30, 0.15, and 0.10 eV, respectively. If the strict alchemy approximation were taken, the value of $\Delta - E_s^{Si}$ would be $E_s^{Ge} - E_s^{Si} = -4.59$ eV in the tight-binding approximation, and a negative E_b^0 ($\simeq -0.10$ eV) would be obtained. When long-range interactions are included, however, the above resonance state becomes a shallow donor level, which is the experimental situation for a P impurity in Si. Our results suggest that V for core excitons in Si is deeper than those implied by alchemy approximations. However, if we use the alchemy approximation as a means of scaling, the value of V for Ge $3p$ core excitons should be deeper than that for Si $2p$ core excitons. Hence, the curve corresponding to $E_b^0 = 0.30$ in Fig. 2 should be a reasonable estimate for Ge $3p$ core-exciton binding energy in alloys. Therefore, we do not expect to see an anomalous behavior of Δ in alloys for this case.

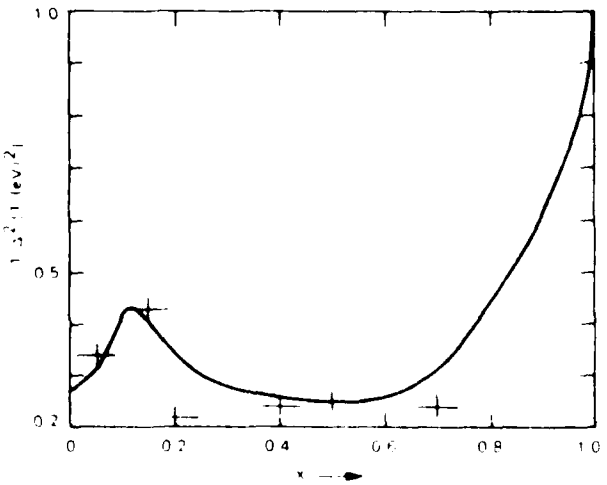


FIG. 3. The calculated $1/\Delta^2$ values (solid line) compared with the experimental results (marks) from Ref. 12. The value of calculated $\Delta(Si)$ is 0.127 eV. The experiment is normalized to the theory at $x = 0.5$.

In summary, the present calculations of the Si $2p$ core-exciton binding energy and linewidth suggest that the exciton level is about 0.15 ± 0.05 eV below the conduction-band edge for pure Si. It follows the X edge for $x > 0.15$ in the Si_xGe_{1-x} alloys, and E_b may eventually reach zero in the dilute limit $x \to 0$. Our value for E_b^0 represents the lower end of the previous measurements,[1-6] but is in good agreement with a recent experimental[13] value of 0.120 ± 0.03 eV. By considering the intrinsic linewidth and the alloy broadening, we can explain the observed relative minimum in the linewidth near $x \simeq 0.15$, without requiring a sudden change of the exciton character. On the basis of this calculation, we expect the corresponding width in pure Si to be even smaller than that observed near $x = 0.15$. We further argue that the binding energy of Ge $3p$ core excitons should be larger than that of Si $2p$ core excitons and there should be no anomaly in the Ge $3p$ linewidth in alloys.

This work was supported in part by the U.S. Department of Defense Advanced Research Projects Agency, under Contract No. MDA 903-83-C-0108 and the Air Force Office of Science Research through Grant No. AFOSR-84-0282. One of us (A.-B.C.) would like to thank Professor W. E. Spicer for his hospitality at Stanford University.

[a]Visiting professor at Stanford University.

[1]W. Eberhardt, G. Kalkoffen, C. Kunz, D. E. Aspnes, and M. Cardona, Phys. Status Solidi (b) **88**, 135 (1978).

[2]M. Altarelli and D. L. Dexter, Phys. Rev. Lett. **29**, 1100 (1972).

[3]F. C. Brown, R. S. Bachrach, and M. Skibowski, Phys. Rev. B **15**, 4781 (1977).

[4]R. S. Bauer, R. S. Bachrach, J. C. McMenamin, and D. E. Aspnes, Nuovo Cimento **39B**, 409 (1977).

[5]G. Margaritondo and J. E. Rowe, Phys. Lett. **59A**, 464 (1977).

[6]G. Margaritondo, A. Franciosi, N. G. Stoffel, and H. S.

Edelman, Solid State Commun. **36**, 298 (1980).
[7]C. Kunz, J. Phys. (Paris), Colloq. **39**, C4-119 (1978).
[8]H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B **24**, 6010 (1981).
[9]M. Altarelli, Phys. Rev. Lett. **46**, 205 (1981).
[10]G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982).
[11]K. E. Newman and J. D. Dow, Solid State Commun. **50**, 587 (1984).
[12]B. A. Bunker, S. L. Hulbert, J. P. Stott, and F. C. Brown, Phys. Rev. Lett. **53**, 2157 (1984).
[13]S. E. Schnatterly and R. D. Carson, Bull. Am. Phys. Soc. **30**, 416 (1985).

PHYSICAL REVIEW B VOLUME 32, NUMBER 6 15 SEPTEMBER 1985

Semiconductor pseudobinary alloys: Bond-length relaxation and mixing enthalpies

A.-B. Chen
Department of Physics, Auburn University, Alabama 36849

A. Sher
SRI International, Menlo Park, California 94025
(Received 1 March 1985)

Harrison's bonding theory, the valence force field (VFF), and an elastic continuum are combined in a study of the substitution energies Δ_s and local (first-shell) bond lengths d_1 of isoelectronic impurities in semiconductors. Explicit expressions for Δ_s and d_1 are derived, which enable us to absorb measured elastic constants into the calculation and to study the chemical effects arising from differences in the covalent radii and polarities. Several models based on VFF alone are also derived for comparison. The full theory and at least five VFF models are found to produce impurity bond lengths in excellent agreement with experiment. The substitution energies are shown to provide good estimates of the mixing enthalpies Ω of pseudobinary alloys and to predict miscibility gaps properly. The chemical shifts in Ω are found to be negative for most cation alloys but positive for anion substitutions.

I. INTRODUCTION

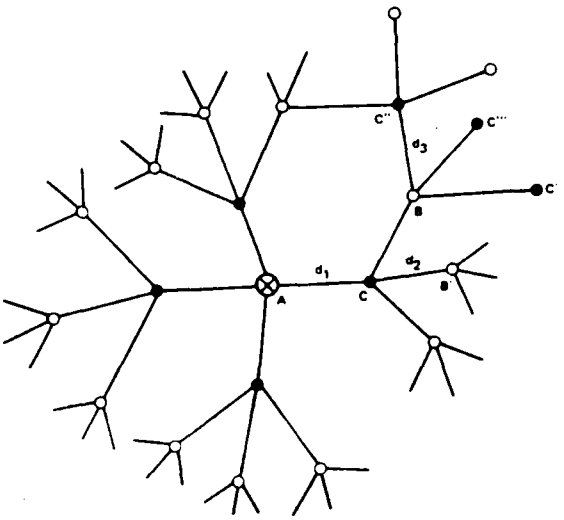
The discovery of a bimodal distribution of the nearest-neighbor bond lengths[1] in $Ga_xIn_{1-x}As$ has sparked considerable interest in the bonding nature of semiconductor alloys.[2–6] This finding has changed the conventional picture of the alloy crystal bond configuration, which has far-reaching implications about the electronic structure, structural stability, and thermodynamics of these materials. Because of the complexity of both the structural and the potential disorder in these alloys, *ab initio* band-structure techniques have not yet evolved to a stage suitable for direct calculations. Therefore, we have extended Harrison's bonding theory[7] to study the alloy structural properties.[5,6] In this paper, we apply an intermediate version of the theory to the dilute-limit case of an isoelectronic impurity.

A particularly useful application of the theory is its perturbation-expansion form, in which measured elastic constants are incorporated to obtain accurate results. This form is also useful for comparison with other previously published models[3,8,9] that are based on the valence-force-field (VFF) (Ref. 10) model alone. Thus, all the factors influencing bond-length relaxation, e.g., strains, boundary conditions, and chemical effects, can be studied. The ability to incorporate the chemical effects is one major difference between this theory and other VFF models.

The remainder of the paper contains the following sections: Sec. II describes a theory for calculating impurity substitution energies. Section III casts the theory into perturbation form and combines it with a valence force field and an elastic continuum. Several VFF models are derived in Sec. IV. The modifications of numerical results due to chemical effects on local bond lengths and alloy mixing enthalpies are summarized and discussed in Sec. V. Conclusions are drawn in the last section, Sec. VI.

II. IMPURITY-SUBSTITUTION ENERGY

Consider the problem of substituting an isoelectronic atom A for a B atom in a zinc-blende compound BC (e.g., In substitutes for Ga in GaAs, as shown in Fig. 1). In



POSITIONS	DISPLACEMENT
C $\frac{a}{\sqrt{3}}(1,1,1)$	$\frac{a}{\sqrt{3}}(\gamma-1,\gamma-1,\gamma-1)$
B $\frac{a}{\sqrt{3}}(0,2,2)$	$\frac{a}{\sqrt{3}}(0,\gamma,\gamma)$
B $\frac{a}{\sqrt{3}}(2,0,2)$	$\frac{a}{\sqrt{3}}(\gamma,0,\gamma)$
C''' $\frac{a}{\sqrt{3}}(1,3,3)$	$\frac{a}{\sqrt{3}}(\gamma,3\gamma-2,3\gamma-2)$
C'''' $\frac{a}{\sqrt{3}}(-1,3,1)$	$\frac{a}{\sqrt{3}}(\gamma-2,3\gamma-2,\gamma)$
C'''' $\frac{a}{\sqrt{3}}(-1,1,3)$	$\frac{a}{\sqrt{3}}(\gamma-2,\gamma,3\gamma-2)$

FIG. 1. A sketch of the flattened picture of a single impurity A in a BC compound. The positions and displacements for those atoms labeled are used in Appendix B.

general, the bond lengths d_1, d_2, d_3, ... for the first-, second-, and third-shell bonds surrounding the impurity are different from the equilibrium values of either the pure BC compound, denoted as d, or the "impurity" compound AC, denoted as $d_I = d(1-\delta_0)$. If A starts being a free atom and B also ends being a free atom, then the energy difference between the final and initial states is defined as the substitution energy and is given by

$$\Delta_s = (E_{\text{def}} + \varepsilon_B) - (E_{\text{pure}} + \varepsilon_A) , \tag{1}$$

where ε_A and ε_B are free-atom energies for A and B, respectively, E_{def} is the total energy of the semiconductor with a defect as sketched in Fig. 1, and E_{pure} is that of the pure host BC crystal. Equation (1) can be written as

$$\begin{aligned}\Delta_s &= (E_{\text{def}} - E_{\text{dis}}) - (\varepsilon_A - \varepsilon_B) + (E_{\text{dis}} - E_{\text{pure}}) \\ &= \Delta_R - (\varepsilon_A - \varepsilon_B) + \Delta_{\text{dis}} ,\end{aligned} \tag{2}$$

where we have added and subtracted a term E_{dis}, which is the total energy of a distorted BC compound with all the atoms held at the positions specified in Fig. 1, except that the central atom is a B atom. Clearly, $\Delta_R = E_{\text{def}} - E_{\text{dis}}$ is a replacement energy, and the distortion energy $\Delta_{\text{dis}} = E_{\text{dis}} - E_{\text{pure}}$ is the energy required to deform a pure BC crystal from its equilibrium lattice to that specified in Fig. 1. Δ_R contains all the chemical terms that arise from different bond lengths and polarities between AC and BC bonds.

Δ_R can be treated most easily by Harrison's bonding theory.[7] In this theory, the energy per bond relative to the vacuum state is

$$E_b = 2\varepsilon_b + V_0 + 6\varepsilon_m^+ + 6\varepsilon_m^- , \tag{3}$$

where ε_b is the energy of the bonding state constructed from the two hybrid orbitals facing each other along the bond direction

$$E_b = \tfrac{1}{2}(\varepsilon_h^A + \varepsilon_h^C) - (V_2^2 + V_3^2)^{1/2} , \tag{4}$$

with ε_h^A and ε_h^C being the energies of the anion and cation hybrid orbitals, respectively. The antibonding energy ε_a has the same form as in Eq. (4), except with a plus sign. V_2 is called the covalent energy, which is the total electronic Hamiltonian matrix element between the two hybrids in question, and the polar energy V_3 is the difference $V_3 = \frac{1}{2}(\varepsilon_h^C - \varepsilon_h^A)$. The ε_m^+ and ε_m^-, the metallization energies, are the shifts of the bonding level caused by interactions with the neighboring antibonding states, where $+$ and $-$ indicate whether the common adjacent atom is a cation or anion. For example, this term for an AC bond labeled by d_1 in Fig. 1 due to an antibonding state labeled by d_2 is given by

$$\varepsilon_{m,C}^{B,A}(2,1) = -\frac{|U_b^A(1)|^2 \, |U_a^B(2)|^2 \, |V_1^C|^2}{\varepsilon_b^A(1) - \varepsilon_a^B(2)} , \tag{5}$$

where A and B denote AC and BC bonds, respectively, $V_1^C = \frac{1}{4}(\varepsilon_s^C - \varepsilon_p^C)$, with ε_s^C and ε_p^C being the s- and p-term values of the common adjacent atom C. $U_b^A(1)$ is the probability amplitude of finding an electron in the hybrid orbital of the C atom in the bonding state of an AC bond with a bond length d_1, whereas $U_a^B(2)$ is the corresponding probability amplitude for the antibonding state of a BC bond of bond length d_2. Finally, V_0 is a repulsive pair potential required to prevent the crystal from collapsing and to guarantee a correct equilibrium bond length. The local perturbation, Eq. (5), is applicable because the square root of the numerator is much smaller than the separation between the antibonding and bonding levels and the valence band is completely filled, so the interaction between the bonding states only spreads the ε_b levels into bands without affecting the center of gravity of the occupied states.

The replacement energy Δ_R of Eq. (2) can now be written explicitly:

$$\begin{aligned}\Delta_R = 4[\, &2\varepsilon_b^A(1) + V_0^A(1) + 6\varepsilon_{m,A}^{A,A}(1,1) + 6\varepsilon_{m,C}^{B,A}(2,1) \\ &- 2\varepsilon_b^B(1) - V_0^B(1) - 6\varepsilon_{m,B}^{B,B}(1,1) - 6\varepsilon_{m,C}^{B,B}(2,1) \\ &+ 6\varepsilon_{mC}^{A,B}(1,2) - 6\varepsilon_{m,C}^{B,B}(1,2)] .\end{aligned} \tag{6}$$

The distortion energy Δ_{dis} of Eq. (2) now involves only BC bonds of different bond lengths. It can be treated with exactly the same procedure for any given set of bond-length distributions. Thus, a straightforward energy minimization procedure can be carried out. The accuracy of this procedure, however, depends in turn on the accuracy of scaling rules for V_2 and V_0 and the input parameters. At present, Harrison's model[7] with $V_2 \propto 1/d^2$ and $V_0 \propto 1/d^4$ and his universal parameters are only semiquantitative. We are improving the quantitative nature of the theory so that the full theory will yield accurate predictions of the structural and thermal properties of semiconductor defects and alloys.

III. PERTURBATION EXPANSION, VALENCE FORCE FIELD, AND ELASTIC MEDIUM

As pointed out earlier, a perturbation expansion of the theory is instructive. This is feasible because the differences $\Delta V_2 = V_2(AC) - V_2(BC)$ and $\Delta V_3 = V_3(AC) - V_3(BC)$ are small compared to each individual value for many of the isoelectronic impurities in III-IV and II-VI compounds. To this end, Eq. (6) is rewritten as

$$\begin{aligned}\tfrac{1}{4}\Delta_R = {} & E_b^A(1) - E_b^B(1) + 6[\varepsilon_{m,C}^{B,A}(2,1) - \varepsilon_{m,C}^{A,A}(1,1)] \\ & - 6[\varepsilon_{m,C}^{A,B}(2,1) - \varepsilon_{m,C}^{B,B}(1,1)] \\ & + 6[\varepsilon_{m,C}^{A,B}(1,2) - \varepsilon_{m,C}^{B,B}(2,2)] \\ & - 6[\varepsilon_{m,C}^{B,B}(1,2) - \varepsilon_{m,C}^{B,B}(2,2)] ,\end{aligned} \tag{7}$$

where $E_b^A(1)$ and $E_b^B(1)$ are energies per bond in Eq. (3) for AC and BC compounds, respectively, with the relaxed bond length $d_1 = d(1-\delta)$. The difference between these energies E_b and the corresponding values at their respective equilibrium bond lengths d_I and d are just the strain energies per bond in uniform deformation:

$$\begin{aligned}E_b^A(1) &= E_b^A(d_I) + 2\sqrt{3}B_I d_I (d_1 - d_I)^2 , \\ E_b^B(1) &= E_b^B(d) + 2\sqrt{3}Bd(d_1 - d)^2 ,\end{aligned} \tag{8}$$

where B_I and B are the bulk moduli for the impurity AC and host BC crystals. The rest of the terms in Eq. (7) are all due to changes in ε_m caused by the differences ΔV_3 and ΔV_2. We shall use Harrison's scaling rules to deduce them.[7] Expanding Eq. (7) to second order in ΔV_3 and $\Delta d = d_2 - d_1$, we write

$$6[\varepsilon^{B,A}_{m,C}(2,1) - \varepsilon^{A,A}_{m,C}(1,1)] = f_I \Delta d - g_I \Delta V_3 + h_I (\Delta d)^2 - W_I \Delta d \, \Delta V_3 + U_I (\Delta V_3)^2 , \tag{9}$$

where f_I, g_I, and so on, are appropriate derivatives evaluated for the impurity crystal AC. When similar expansions are made for the rest of the terms in Eq. (7), it becomes [with $d_1 = d(1-\delta)$, $d_I = d(1-\delta_0)$]

$$\tfrac{1}{4}\Delta_R = \Delta E_b + (f_I - f)\Delta d - (g_I - g)\Delta V_3 + (h_I + h)(\Delta d)^2 - (W_I + W)\Delta d \, \Delta V_3 + (U_I + U)(\Delta V_3)^2 + 2\sqrt{3} B_I d^3 (\delta - \delta_0)^2 - 2\sqrt{3} B d^3 \delta^2 , \tag{10}$$

where

$$\Delta E_b = E_b^A(d_I) - E_b^B(d) - \tfrac{1}{4}(\varepsilon_A - \varepsilon_B) \tag{11}$$

is just the difference in the binding energy per bond between the "BC" and "AC" crystals. In Eq. (10) the coefficients f, g without a subscript are those for the host BC system. It is convenient to define an excess energy $\Delta E = \Delta_s/4 - \Delta E_b$, which is the extra energy per bond required for the impurity substitution over and above the binding-energy difference between the BC and AC crystals. The binding-energy difference accounts for much of the substitution energy; however, the correction measured by the excess energy can be significant. The excess energy results from strain energies and chemically driven charge redistributions around the defect. Using Eqs. (2) and (10) and defining $F = f_I - f$ and $G = g_I - g$, we can write ΔE up to second order in ΔV_3 and Δd as

$$\Delta E = 2\sqrt{3} B_I d^3 (\delta - \delta_0)^2 - 2\sqrt{3} B d^3 \delta^2 + F \Delta d - G \Delta V_3 + H(\Delta d)^2 + W \Delta d \, \Delta V_3 + U(\Delta V_3)^2 + \tfrac{1}{4}\Delta_{\rm dis} , \tag{12}$$

where $H = h_I + h$, $W = w_I + w$, and $U = u_I + u$.

To treat the distortion energy $\Delta_{\rm dis}$, we divide the crystal into two regions. Inside a sphere of some radius R measured from the impurity, the strain energy is taken to be the valence-force-field[10] value:

$$\Delta_{\rm dis}^{\rm (in)} = \frac{3}{8d^2}\sum_i \alpha_i [\Delta(\mathbf{d}_i \cdot \mathbf{d}_i)]^2 + \frac{3}{8d^2}\sum_i{}' \sum_{j\,(j<i)} \beta_{ij} [\Delta(\mathbf{d}_i \cdot \mathbf{d}_j)]^2 , \tag{13}$$

where i sums over all the bonds inside R and the pairs in the β terms include those that have adjacent atoms inside R and on the boundary. The parameters α and β are force constants to be considered later. $\Delta(\mathbf{d}\cdot\mathbf{d}_j) = \mathbf{d}_i \mathbf{d}_j - \mathbf{d}_i^{(0)} \cdot \mathbf{d}_j^{(0)}$ measures the change of the dot product between bond vectors due to distortions. Outside R we assume an elastic continuum with radial displacements inversely proportional to the square of the distance from the center. The elastic energy in this medium can be shown to be (see Appendix A)

$$\Delta_{\rm dis}^{\rm (out)} = RCu^2 , \tag{14}$$

where the effective shear coefficient is

$$C = 4\pi[0.4(C_{11} - C_{12}) + 1.2C_{44}]$$

and u is the displacement at R. In view of the fact that the bonds d_1 and d_2 are coupled through the chemical terms in Eq. (10), the smallest logical radius R is the second-shell atomic distance, namely $R = 2\sqrt{2}d/\sqrt{3}$. Atoms on this boundary have displacements of the forms $\mathbf{u} = d(\gamma,\gamma,0)/\sqrt{3}, \ldots$. Thus, $u = \sqrt{2}\gamma d/\sqrt{3}$ and the elastic energy in the continuum is

$$\Delta_{\rm dis}^{\rm (out)} = \tfrac{4}{3}\sqrt{2/3} C \gamma^2 d^3 . \tag{15}$$

The distortion energy represented by Eq. (14) contains six different contributions (see Appendix B): the bond-stretching energy of the four first-shell bonds $6\alpha\delta^2 d^2$, the β terms from the first-shell bonds, $\beta\delta^2 d^2$, the α terms from the second-shell bonds, $2\alpha(\delta + 2\gamma)^2 d^2$, the β terms between the first- and second-shell bonds, $2\beta(\delta+\gamma)^2 d^2$, the β terms among the second-shell bonds, $2\beta\delta^2 d^2$, and finally the β terms between the second-shell bonds and those in the continuum,

$$\tfrac{1}{2}\beta d^2[(3\delta + \lambda_1\gamma)^2 + (\delta + \lambda_2\gamma)^2] ,$$

where $\lambda_1 = 40\sqrt{2}/(19\sqrt{19})$ and $\lambda_2 = 2 - 40\sqrt{2}/(11\sqrt{11})$.

To assemble all the contributions to Eq. (12), we need to consider the assignments of the elastic constants and the force constants α and β in VFF. While the experimental values[11] of C_{11}, C_{12}, and C_{44} can be used for the elastic constants, α and β have to be deduced. If Martin's original procedure[10] (also followed by Martins and Zunger[3]) is used, then Eq. (13) alone will not produce the correct (experimental) bulk moduli. There are small corrections due to Madelung terms, which are hard to treat in the case of nonuniform distortions. A simpler procedure is adopted here. We use the experimental bulk moduli for B_I and B in Eq. (12) and experimental elastic constants to calculate C of Eq. (14) and then force α and β in the VFF to produce the correct bulk moduli B and shear coefficients $C_{11} - C_{12}$. Such an approach is also consistent with Harrison's bonding theory[7] and other approaches in which the Coulomb forces are automatically incorporated in the band and bond energies, and do not need to be redundantly treated. With our procedure, the bulk modulus is simply related to the force constants by $B = (3\alpha + \beta)/(4\sqrt{3}d)$. Table I lists our α and β values. We want to point out in advance, however, that the numerical results deduced from our sets and those of Martin of α and β do not introduce differences more than the present experimental uncertainties in the local bond length (~ 0.01 Å) and the mixing enthalpies (> 0.5 kcal/mole).

Using the above procedure and adding all contributions, the excess energy per bond from Eqs. (10) and (12) is the full perturbation theory (FPT) result

TABLE I. The bond lengths d (in Å), valence force constants α and β (N/m), shear coefficients C of the continuum (in 10^{11} ergs/cm^3), melting temperatures T_m (K), and Liedermann ratios χ_m for the compounds used in this paper.

Compound	d	α	β	C	T_m[a]	χ_m
AlP	2.367	44.323	8.068	122.396	1773	0.070
AlAs	2.451	40.849	8.717	112.695	1873	0.073
AlSb	2.656	34.073	6.900	85.351	1323	0.062
GaP	2.360	44.764	10.737	145.921	1510	0.064
GaAs	2.448	39.235	9.159	121.844	1738	0.071
GaSb	2.640	31.876	7.347	89.372	985	0.055
InP	2.541	40.363	6.543	91.785	1343	0.059
InAs	2.622	33.203	5.752	78.816	1215	0.061
InSb	2.805	28.557	4.891	60.721	798	0.049
ZnS	2.342	40.429	5.273	89.272	2123	0.081
ZnSe	2.454	32.200	4.562	82.687	1788	0.080
ZnTe	2.637	29.445	4.659	62.430	1511	0.071
CdTe	2.806	26.569	2.722	38.453	1371	0.067
HgTe	2.798	26.396	2.746	40.363	943	0.056

[a]Reference 30.

$$\Delta E = [\,3\alpha_I(\delta-\delta_0)^2/2+\beta_I(\delta-\delta_0)^2/2+\alpha(\delta+2\gamma)^2/2 + \beta(\delta+\gamma)^2/2+\beta\delta^2/4+\beta(3\delta+\lambda_1\gamma)^2/8 + \beta(\delta+\lambda_2\gamma)^2/4]d^2+\sqrt{2}C\gamma^2d^3/(3\sqrt{3})+\Delta E_{\rm ch}\,, \tag{16}$$

where the chemical contribution is written as

$$\Delta E_{\rm ch}=F_{\rm ch}\,\Delta d+H(\Delta d)^2+\Delta E_p = \tfrac{2}{3}F_{\rm ch}(2\delta+\gamma)d+\tfrac{4}{9}H(2\delta+\gamma)^2d^2 + [\,U(\Delta V_3)^2-G\,\Delta V_3]\,, \tag{17}$$

where $\Delta F_{\rm ch}=F-W\,\Delta V_3$ and $\Delta E_p=U(\Delta V_3)^2-G\,\Delta V_3$. $F_{\rm ch}$ is a chemical force, which when it is positive tends to push the C atom away from the impurity atom A. This force arises from the difference in the bond tensions induced between the AC and BC bonds adjacent to C because the neighboring antibinding states are different from those of their respective host states. ΔE_p is due to the difference in the polarities ΔV_3 alone and is independent of the displacement. Finally, H can be regarded as a chemically induced force constant, which when it is positive tends to restrain the lattice from distortion and increases the elastic energy.

The equilibrium requirements $\partial(\Delta E)/\partial\delta=0$ and $\partial(\Delta E)/\partial\gamma=0$ then lead to the solution $\gamma=Q\delta$, and δ is given by

$$\delta=(\delta_0+\delta_0')/\{1+[\,\alpha(1-2Q)+\beta(17/4-\lambda Q) + 16H(1-2Q)/9]/(3\alpha_I+\beta_I)\}\,, \tag{18}$$

where the constant λ is $1+3\lambda_1/4+\lambda_2/2$, and

$$\delta_0'=-4F_{\rm ch}(1-Q/2)/[3d(3\alpha_I+\beta_I)]\,, \tag{19}$$

with $Q=2J/K$, $J=\alpha+\lambda\beta/2+8H/9$, and

$$K=4\alpha+2\sqrt{2}Cd/(3\sqrt{3})+(1+\lambda_1^2/4+\lambda_2^2/2)\beta+8H/9\,.$$

IV. VALENCE-FORCE-FIELD MODELS

In this section we consider several models based on the valence force field. These models have been used frequently to explain the impurity bond relaxation.[3,4,8,9] We shall first derive the explicit expressions for these models and then connect them with the existing results.

A. Model A: Third-shell atoms and beyond are fixed at their pure crystal positions

Let the bond lengths surrounding the impurity again be $d_1=d(1-\delta)$ and let the second-shell atoms have radial displacements of the forms $(d/\sqrt{3})$, $(\gamma,\gamma,0)$, etc. Beyond and including the third shell, all the other atoms are held at their pure-crystal positions. There are nine different contributions to the strain energy in VFF (see Appendix B): the α terms from the four bonds surrounding the impurity, $6\alpha_I(\delta-\delta_0)^2d^2$; the β terms among the six pairs of these bonds, $\beta_I(\delta-\delta_0)^2d^2$; the β terms between the first- and second-shell bonds, $2\beta(\delta+\gamma)^2d^2$; the α terms from the second-shell bonds, $2\alpha(\delta+2\gamma)^2d^2$; the β terms among the second-shell bonds, $2\beta\delta^2d^2$; the β terms between the second- and third-shell bonds, $\frac{9}{2}\beta\delta^2d^2+\beta(\delta+2\gamma)^2d^2$; the α terms from the third-shell bonds, $8\alpha\gamma^2d^2$; the β terms among the third-shell bonds, $4\beta\gamma^2d^2$; and the β terms between the third- and fourth-shell bonds, $6\beta\gamma^2d^2$. Thus,

the excess energy (in this case $\frac{1}{4}$ times the strain energy) becomes

$$\Delta E=[\tfrac{3}{2}\alpha_I(\delta-\delta_0)^2+\tfrac{1}{4}\beta_I(\delta-\delta_0)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{13}{8}\beta\delta^2+\tfrac{1}{4}\beta(\delta+2\gamma)^2+2\alpha\gamma^2+\tfrac{3}{2}\beta\gamma^2]d^2 . \quad (20)$$

The minimization of ΔE with respect to δ and γ leads Eq. (20) to $\gamma=-\delta/4$, and

$$\delta=\delta_0/[1+(\alpha+17\beta/2)/(6\alpha_I+\beta_I)] . \quad (21)$$

We note that there is some ambiguity in the third contribution listed above for the β terms between the impurity and host bonds. The value of β could be chosen as one of these combinations β, β_I, $\frac{1}{2}(\beta+\beta_I)$, $\sqrt{\beta\beta_I}$, or other proper combinations. Because the values of β and β_I are comparable and β values are much smaller than α (see Table I), the results for δ and ΔE are not too sensitive to the choice. There is also some ambiguity in the values for $\mathbf{d}_i^{(0)}\cdot\mathbf{d}_j^{(0)}$ for the "undistorted" crystal. The $-d^2/3$ used is the simplest choice. A different choice will not affect the results for δ at all, but will make ΔE slightly different. In fact, model A was first used by Martins and Zunger.[3] However, their expression for δ is different from Eq. (21) because they made different choices of the two quantities just mentioned. Nevertheless, Sec. IV will show that these two expressions yield very similar results. These ambiguities do not occur in the full theory in Sec. III, where the impurity-host interactions are taken into account naturally by the replacement energy Δ_R [see Eq. (10)].

B. Model B: Second-shell atoms connect to a fixed boundary

This model corresponds to $\gamma=0$ in model A. So we have

$$\Delta E=[\tfrac{3}{2}\alpha_I(\delta-\delta_0)^2+\tfrac{1}{4}\beta_I(\delta-\delta_0)^2+\tfrac{1}{2}\alpha\delta^2+\tfrac{19}{8}\beta\delta^2]d^2 \quad (22)$$

and

$$\delta=\delta_0/[1+(\alpha+19\beta/4)/(3\alpha_I+\tfrac{1}{2}\beta_I)] . \quad (23)$$

This expression will be used to study the effect of truncation.

C. Model C: Simple spring model

If all the β's in Eqs. (22) and (23) are set equal to zero, we have the simple spring model with

$$\Delta E=[\tfrac{3}{2}\alpha_I(\delta-\delta_0)^2+\tfrac{1}{2}\alpha\delta^2]d^2 \quad (24)$$

and

$$\delta=\delta_0/(1+\tfrac{1}{3}\alpha/\alpha_I) . \quad (25)$$

The spring model recently discussed by Shih *et al.*[8] corresponds to Eq. (25) with $\alpha=\alpha_I$, so $\delta/\delta_0=\frac{3}{4}$.

D. Model D: VFF with the continuum connected to the second-shell atoms

1. Model D1

In this case, ΔE only contains the first five contributions listed for case A plus the elastic energy in the continuum. However, the β terms between the second- and third-shell bonds are modified because atoms outside R in the continuum now have radial displacements proportional to the inverse of the square of the radius. The result is

$$\Delta E=\left[\tfrac{3}{2}\alpha_I(\delta-\delta_0)^2+\tfrac{1}{4}\beta_I(\delta-\delta_0)^2+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2+\tfrac{1}{2}\beta\delta^2+\tfrac{1}{8}\beta(3\delta+\lambda_1\gamma)^2+\tfrac{1}{4}\beta(\delta+\lambda_2\gamma)^2+\frac{\sqrt{2}}{3\sqrt{3}}Cd\gamma^2\right]d^2 , \quad (26)$$

where λ_1 and λ_2 are the same as the constants that appear in Eq. (16). The corresponding equilibrium condition can be shown to be

$$\delta=\delta_0/\{1+[\alpha(1-2Q)+19\beta/4-\beta(1+3\lambda_1/4+\lambda_2/4)Q]/(3\alpha_I+\beta_I/2)\} , \quad (27)$$

where

$$Q=\frac{(2\alpha+\beta+3\lambda_1\beta/4+\lambda_2\beta/2)}{(4\alpha+2\sqrt{2}Cd/3\sqrt{3}+\beta+\tfrac{1}{4}\lambda_1^2\beta+\tfrac{1}{2}\lambda_2^2\beta)} . \quad (28)$$

2. Model D2

A comparison between Eq. (26) and the full perturbation theory, Eq. (16), shows two major differences. First, all the chemical terms are absent in Eq. (26). Secondly, the terms

$$\tfrac{1}{2}\beta_I(\delta-\delta_0)^2d^2+\tfrac{1}{4}\beta\delta^2d^2$$

in Eq. (16) become

$$\tfrac{1}{4}\beta_I(\delta-\delta_0)^2d^2+\tfrac{1}{2}\beta\delta^2d^2$$

in Eq. (26). This difference in the strain energy will mask the true effects of chemical forces if δ from Eq. (26) is compared with FPT. A better way to study the chemical effects is to use the following equation:

$$\Delta E=[\tfrac{1}{2}\alpha_I(\delta-\delta)^2+\tfrac{1}{2}\beta_I(\delta-\delta_0)^2+\tfrac{1}{2}\alpha(\delta+2\gamma)^2+\tfrac{1}{2}\beta(\delta+\gamma)^2+\tfrac{1}{4}\beta\delta^2+\tfrac{1}{8}\beta(3\delta+\lambda_1\gamma)^2+\tfrac{1}{4}\beta(\delta+\lambda_2\gamma)^2]d^2+\frac{\sqrt{2}}{3\sqrt{3}}C\gamma^2d^3\,, \tag{29}$$

which is Eq. (16) with all the chemical terms neglected. The corresponding δ becomes

$$\delta=\delta_0/\{1+[\alpha(1-2Q)+\tfrac{17}{4}\beta-\beta(1+\tfrac{3}{4}\lambda_1+\tfrac{1}{4}\lambda_2)Q]/(3\alpha_I+\beta_I)\}\,, \tag{30}$$

with Q still given by Eq. (28).

E. Model E: Continuum connected to the first-shell atoms

In this case, $\gamma=-3\sqrt{3}\delta/(8\sqrt{2})$ and ΔE only includes the first three contributions listed in model A plus the strain energy of the continuum:

$$\Delta E=\left\{\tfrac{1}{2}\alpha_I(\delta-\delta_0)^2+\tfrac{1}{4}\beta_I(\delta-\delta_0)^2+\tfrac{1}{2}\beta\left[1-\frac{3\sqrt{3}}{8\sqrt{2}}\right]^2\gamma^2+\tfrac{1}{4}Cd\delta^2\right\}d^2\,. \tag{31}$$

The relaxation parameter is given by

$$\delta=\delta_0\Big/\left\{1+\left[\tfrac{1}{2}Cd+\left[1-\frac{3\sqrt{3}}{8\sqrt{2}}\right]^2\beta\right]\Big/(3\alpha_I+\tfrac{1}{2}\beta_I)\right\}\,. \tag{32}$$

We note that the continuum model used to estimate the bond-length relaxation by Baldereschi and Hopfield[9] corresponds to Eq. (32) without the β terms, which yields $\delta/\delta_0\approx 0.4$ to 0.5, rather than the proper values around 0.7 to 0.8.

V. ALLOY MIXING ENTHALPY

The impurity-substitution excess energies ΔE provide a first estimate of the mixing enthalpies of pseudobinary alloys. Most current thermodynamics theories of semiconductor alloys are based on an extension of the binary solution model.[12] In this model, the mixing Helmholtz energy of an $A_xB_{1-x}C$ alloy is defined as

$$\Delta F_m=F_{\text{alloy}}-(xF_{AC}+yF_{BC})\,, \tag{33}$$

where $y=1-x$, and F_{AC} and F_{BC} are the respective free energies of the pure AC and BC compounds at the same temperature. Because the C atoms occupy a sublattice, the nearest neighbors of A and B atoms in the alloy are the C atoms. Thus, the pair potentials that enter the binary solution theory are now the second-neighbor interactions. Let N_{AA}, N_{AB}, and N_{BB} be, respectively, the numbers of the second-neighbor AA, AB, and BB pairs, with corresponding pair interaction energies ϵ_{AA}, ϵ_{AB}, and ϵ_{BB}. For tetrahedral semiconductors, there are a total of $6N$ second-neighbor pairs for a crystal containing N unit cells. Denote the ratios N_{AA}, N_{AB}, and N_{BB} to $6N$ as r_{AA}, $r_{AB}=r$, and r_{BB}, respectively. Then those ratios are related to the alloy composition by $r_{AA}=x-r/2$ and $r_{BB}=y-r/2$. The mixing free energy has two terms,

$$\Delta F_m=\Delta E_m-T\Delta S\,, \tag{34}$$

where the mixing energy is given by

$$\begin{aligned}\Delta E_m&=E_{\text{alloy}}-(xE_{AC}+yE_{BC})\\&=6N(\epsilon_{AA}r_{AA}+\epsilon_{AB}r_{AB}+\epsilon_{BB}r_{BB})-6N(x\epsilon_{AA}+y\epsilon_{BB})\\&=6Nr\,\Delta\epsilon\,,\end{aligned} \tag{35}$$

where

$$\Delta\epsilon=\epsilon_{AB}-\tfrac{1}{2}(\epsilon_{AA}+\epsilon_{BB})\,. \tag{36}$$

The mixing entropy ΔS can also be written from a simple generalization of the random distribution.[12] For modest pressure, ΔE is the same as the mixing enthalpy ΔH_m.

Now the pair interaction energies can be approximately related to the impurity-substitution energies by

$$\Delta_s(A\text{ in }BC)\approx 12(\epsilon_{AB}-\epsilon_{BB}) \tag{37}$$

and

$$\Delta_s(A\text{ in }BC)\approx 12(\epsilon_{AB}-\epsilon_{AA})\,. \tag{38}$$

Thus, $\Delta\epsilon$ of Eq. (36) becomes

$$\begin{aligned}\Delta\epsilon&=\tfrac{1}{24}[\Delta_s(A\text{ in }BC)+\Delta_s(B\text{ in }AC)]\\&=\tfrac{1}{6}[\Delta E(A\text{ in }BC)+\Delta E(B\text{ in }AC)]\,.\end{aligned} \tag{39}$$

Usually, the experimental ΔH_m is written as

$$\Delta H_m=x(1-x)\Omega\,, \tag{40}$$

which is equivalent to assuming a random distribution, i.e., $r=2x(1-x)$. Using this expression for r and comparing Eqs. (40) and (35), we see that the mixing enthalpy parameter Ω is given by

$$\Omega=2[\Delta E(A\text{ in }BC)+\Delta E(B\text{ in }AC)]\,. \tag{41}$$

This connection provides a further check of the theory.

VI. NUMERICAL RESULTS AND DISCUSSION

A. Chemical terms

Table II lists $\delta_0=1-d_I/d$, $\delta=1-d_1/d$, the excess energy (per bond) ΔE for the full theory and its corresponding VFF model D2, and the terms derived from the metallization energies, δ_0' [Eq. (19)], F_{ch}, H, ΔE_p, and ΔE_{ch} [Eq. (17)]. The appropriate derivatives f, g, h,... [see Eq. (9)] are computed using the atomic-term values that we have generated from impurity-level[13] and structural studies.[5]

For substitutions involving the cation pair (Ga,Al), F_{ch} has the same sign as δ_0, which means that F_{ch} prevents relaxation and thus tends to increase the strain energy. The chemical forces H are also significant. As a result,

TABLE II. Comparison between the full theory and the corresponding VFF model D2 to study the effects of chemical terms. All ΔE's are in units of kcal/mole band.

		Model D2					Full theory				
Impurity	Host	δ_0	δ	ΔE	δ_0'	δ	F_{ch} (10^{-10}N)	H (N/m)	ΔE_p	ΔE_{ch}	ΔE
Ga	AlP	0.003	0.002	0.001	−0.001	0.001	0.150	4.581	−0.016	−0.016	−0.013
Al	GaP	−0.003	−0.002	0.001	0.001	−0.001	−0.150	4.581	−0.016	−0.016	−0.013
Ga	AlAs	0.001	0.001	0.000	−0.002	−0.001	0.243	5.733	−0.020	−0.021	−0.018
Al	GaAs	−0.001	−0.001	0.000	0.002	0.001	−0.243	5.733	−0.020	−0.021	−0.018
Ga	AlSb	0.006	0.004	0.005	−0.004	0.001	0.389	5.632	−0.054	−0.053	−0.039
Al	GaSb	−0.006	−0.004	0.005	0.004	−0.001	−0.389	5.632	−0.054	−0.053	−0.039
In	GaP	−0.077	−0.052	0.959	−0.006	−0.054	0.699	3.778	−0.188	−0.219	0.742
Ga	InP	0.071	0.056	0.734	0.005	0.057	−0.699	3.778	−0.188	−0.206	0.530
In	GaAs	−0.071	−0.048	0.752	−0.009	−0.050	0.804	4.778	−0.257	−0.283	0.472
Ga	InAs	0.066	0.052	0.592	0.007	0.054	−0.804	4.778	−0.257	−0.265	0.330
In	GaSb	−0.062	−0.043	0.554	−0.004	−0.042	0.352	5.201	−0.363	−0.308	0.247
Ga	InSb	0.059	0.046	0.445	0.004	0.044	−0.352	5.201	−0.363	−0.287	0.160
In	AlP	−0.074	−0.053	0.761	−0.007	−0.056	0.769	3.506	−0.035	−0.087	0.679
Al	InP	0.068	0.053	0.674	0.006	0.056	−0.769	3.506	−0.035	−0.083	0.596
In	AlAs	−0.070	−0.048	0.705	−0.010	−0.051	0.942	4.437	−0.048	−0.111	0.602
Al	InAs	0.065	0.052	0.576	0.008	0.054	−0.942	4.437	−0.048	−0.099	0.485
In	AlSb	−0.056	−0.039	0.440	−0.008	−0.041	0.689	4.979	−0.061	−0.073	0.369
Al	InSb	0.053	0.042	0.368	0.007	0.044	−0.689	4.979	−0.061	−0.061	0.310
Cd	ZnTe	−0.064	−0.048	0.432	−0.003	−0.050	0.202	−0.484	−0.005	−0.064	0.373
Zn	CdTe	0.060	0.050	0.314	0.002	−0.053	−0.202	−0.484	−0.005	−0.072	0.247
Hg	CdTe	0.003	0.002	0.001	0.004	0.005	−0.278	−0.753	−0.018	−0.026	−0.018
Cd	HgTe	−0.003	−0.002	0.001	−0.004	−0.005	0.278	−0.753	−0.018	−0.026	−0.018
Hg	ZnTe	−0.061	−0.045	0.392	−0.001	0.046	0.075	0.002	0.052	0.037	0.429
Zn	HgTe	0.058	0.048	0.286	0.001	0.049	−0.075	0.002	0.052	0.035	0.322
As	AlP	−0.035	−0.026	0.179	0.001	−0.025	−0.085	0.717	−0.005	0.008	0.187
P	AlAs	0.034	0.025	0.185	−0.001	0.025	0.085	0.717	−0.005	0.008	0.194
As	GaP	−0.037	−0.025	0.226	0.002	−0.024	−0.181	1.078	−0.011	0.012	0.240
P	GaAs	0.036	0.027	0.211	−0.001	0.025	0.181	1.078	−0.011	0.014	0.228
As	InP	−0.032	−0.023	0.136	0.001	−0.022	−0.057	0.919	−0.003	0.008	0.144
P	InAs	0.031	0.024	0.128	−0.001	0.024	0.057	0.919	−0.003	0.009	0.138
Sb	AlAs	−0.084	−0.058	1.024	0.008	−0.051	−0.815	0.644	−0.180	0.002	1.060
As	AlSb	0.077	0.059	0.919	−0.007	0.053	0.815	0.644	−0.180	0.027	0.984
Sb	GaAs	−0.078	−0.052	0.908	0.018	−0.040	−1.599	0.927	−0.363	−0.106	0.929
As	GaSb	0.073	0.055	0.823	−0.014	0.044	1.599	0.927	−0.363	−0.061	0.904
Sb	InAs	−0.070	−0.051	0.603	0.010	−0.042	−0.824	0.855	0.171	−0.009	0.645
As	InSs	0.065	0.051	0.551	−0.009	0.044	0.824	0.855	−0.171	0.008	0.613
Sb	AlP	−0.122	−0.085	2.007	0.010	−0.077	−0.944	0.645	0.241	0.074	2.127
P	AlSb	0.109	0.085	1.855	−0.008	0.078	0.944	0.645	−0.241	0.123	2.030
Sb	GaP	−0.119	−0.075	2.132	0.021	−0.061	−1.868	0.930	−0.505	−0.046	2.244
P	GaSb	0.106	0.083	1.806	−0.015	0.070	1.868	0.930	−0.505	0.093	2.084
Sb	InP	−0.104	−0.072	1.383	0.011	−0.063	−0.922	0.854	−0.214	0.059	1.501
P	InSb	0.094	0.077	1.191	−0.008	0.069	0.922	0.854	−0.214	0.123	1.379
Se	ZnS	−0.048	−0.036	0.231	−0.001	−0.036	0.077	0.645	0.003	0.000	0.231
S	ZnSe	0.046	0.037	0.221	0.001	0.037	−0.077	0.645	0.003	0.001	0.222
Te	ZnSe	−0.075	−0.056	0.550	0.000	−0.056	0.028	0.635	0.008	0.024	0.574
Se	ZnTe	0.069	0.054	0.532	0.000	0.054	−0.028	0.635	0.008	0.025	0.557
Te	ZnS	−0.126	−0.092	1.565	−0.001	−0.092	0.101	0.644	0.022	0.041	1.606
S	ZnTe	0.112	0.091	1.446	0.001	0.091	−0.101	0.644	0.022	0.051	1.496

all six cases involving this pair have nearly equal d_1 and d_2, i.e., the small bond-length differences are made even smaller in the alloy. The excess energies all become negative, mainly because ΔE_p is negative. For the systems involving the (Ga,In) and (In,Al) pairs, F_{ch} has the opposite sign from δ_0, so δ_0' and δ_0 have the same sign. The chemical force favors bond distortion. However, because H is positive and introduces an increase in the denominator of Eq. (18), most of the effect of δ_0' is cancelled. For cases involving (Ga,In), the polarity contributions ΔE_p are all negative. The $F_{ch}\,\Delta d$ term is negative, but $H(\Delta d)^2$ is positive, so they cancel to a certain degree and leave ΔE lowered primarily because of ΔE_p. While ΔE_p is still negative for the (In,Al) substitutions, its magnitude is reduced considerably. The other chemical energies $F_{ch}\,\Delta d + H(\Delta d)^2$ can be as large as ΔE_p, but the overall reductions of ΔE are only a fraction of those for the (Ga,In) cases. For the several II-VI systems studied, both F_{ch} and H are small and the net changes in δ have the same sign as δ_0. However, because δ_0 is small in the (Cd,Hg) substitutions, F_{ch} actually causes a reversal in which the short bond length becomes shorter and the longer one becomes longer. This is the only exceptional case of this type found for all the systems studied. The change of ΔE due to chemical terms in the (Hg,Zn) substitution is also peculiar—it increases mainly because ΔE_p is positive.

Next, we examine the anion substitutions. For the groups involving the (P,As) pair, the chemical shifts are all small, but the trend is less toward relaxation and larger ΔE. The groups involving (As,Sb) and (P,Sb) pairs behave very similarly: F_{ch} are significant and are opposing relaxations, i.e., δ_0' and δ_0 have opposite signs. At the same time, the H values are several times smaller than those for the corresponding III-V cation substitution case. Thus, most of δ_0' translates into a real reduction of the ratio δ/δ_0 and consequently introduces extra strain energy. Although the ΔE_p energies are significant and negative, $F_{ch}\,\Delta d$ are positive and the net ΔE_{ch} can be either positive or negative. However, the induced-strain energy due to reduction of the δ/δ_0 makes all ΔE positive for these two groups of systems. For II-VI systems, all the chemical effects again are small, but the net chemical changes on ΔE are slightly positive.

The above discussion can be summarized in Fig. 2, where the excess energies ΔE calculated from the full perturbation theory and Model D2 are plotted normalized to the results of the simple spring model of Shih, Spicer, Harrison, and Sher (SSHS) (Ref. 8), i.e., Eq. (24) with $\alpha_1=\alpha$, so $\Delta E/(\frac{3}{8}\alpha d^2)=\delta_0^2$. The calculated ΔE rises faster for $\delta_0 \gtrsim 0$ than for $\delta_0 \lesssim 0$, mainly because $\alpha/\alpha_1=1$. In fact, if the relation[7,10] $\alpha/\alpha_1=(d_1/d)^S$ with S of order of 3 to 5 is used in Eq. (24), we obtain a percentage correction of $S\delta_0/4$ to the SSHS results, which explains the skewed behavior of the curve. It is also clear that ΔE rises faster than δ_0^2 for larger δ_0. However, the zeroth-order theory of SSHS is clearly an excellent representation of the global features of ΔE. The results from model D2 are closer to the parabolic form than those from FPT. The figure clearly shows the general trends; the chemical terms cause negative shifts in ΔE for cation substitutions and positive shifts for anion impurities. It is also clear that the chemical shifts can be very large. These effects will have important consequences on the alloy mixing enthalpies to be discussed later.

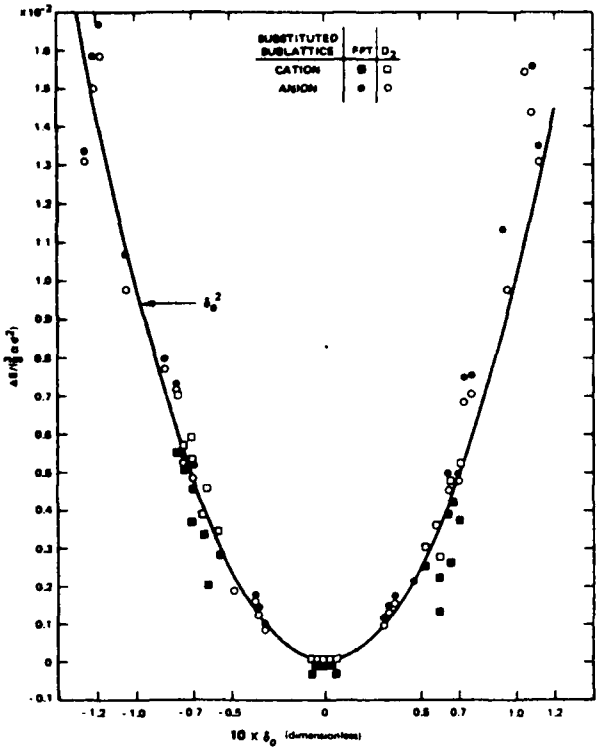


FIG. 2. The excess energies ΔE over $\frac{3}{8}\alpha d^2$ calculated from the full perturbation theory (FPT) and its corresponding valence-force-field model D2. The solid curve corresponds to Eq. (24) with $\alpha=\alpha_1$.

B. Impurity bond length

Table III lists the impurity bond lengths d_1 calculated from different models, while a comparison of theory and the available experimental data[1,14,15] is presented in Table IV. The actual size of changes in d_1 induced by the chemical terms can be seen by comparing model D2 with the full theory. Except for the systems involving the substitution of (Ga,As) and (P,Sb) pairs (where changes range from 0.01 to 0.03 Å), all the chemically induced changes are less than 0.01 Å. Comparison among models A, B, and C shows that, while extending the boundary helps the relaxation (compare model B to A), i.e., δ/δ_0 is closer to 1, the inclusion of the bond-bending forces (the β terms) (compare model B with C) prevents it. The simple spring model (model C), which contains neither of these terms, evidently represents a delicate cancellation of these effects and predicts results close to those of the full perturbation theory and experiment. Although the d_1 values of model C are often very close to those of model A, there are cases [e.g., Ga(P,Sb)] in which model C can differ from model

TABLE III. Calculated impurity local bond lengths (in Å) from the full theory and several valence force models discussed in Sec. III, and their comparison with the values calculated by Martins and Zunger (Ref. 3).

Impurity	Host	A	B	C	D1	D2	E	FPT	MZ
Ga	AlP	2.362	2.363	2.362	2.362	2.362	2.364	2.364	
Al	GaP	2.365	2.364	2.365	2.365	2.365	2.365	2.363	
Ga	AlAs	2.449	2.449	2.449	2.449	2.449	2.450	2.452	
Al	GaAs	2.450	2.450	2.450	2.450	2.450	2.450	2.447	
Ga	AlSb	2.645	2.646	2.644	2.645	2.644	2.649	2.653	
Al	GaSb	2.651	2.650	2.652	2.652	2.652	2.650	2.644	
In	GaP	2.477	2.462	2.492	2.479	2.483	2.435	2.487	2.474
Ga	InP	2.406	2.421	2.402	2.395	2.399	2.409	2.396	2.409
In	GaAs	2.559	2.544	2.573	2.561	2.565	2.518	2.570	2.556
Ga	InAs	2.492	2.506	2.486	2.482	2.485	2.496	2.481	2.495
In	GaSb	2.747	2.734	2.760	2.749	2.754	2.710	2.750	2.739
Ga	InSb	2.683	2.697	2.678	2.673	2.676	2.686	2.680	2.683
In	AlP	2.487	2.472	2.494	2.490	2.493	2.447	2.498	2.480
Al	InP	2.412	2.427	2.408	2.401	2.405	2.415	2.400	2.414
In	AlAs	2.561	2.546	2.572	2.563	2.568	2.523	2.575	2.553
Al	InAs	2.493	2.506	2.487	2.483	2.487	2.497	2.480	2.495
In	AlSb	2.754	2.741	2.763	2.756	2.760	2.721	2.765	2.746
Al	InSb	2.693	2.705	2.689	2.685	2.687	2.696	2.683	2.693
Cd	ZnTe	2.756	2.740	2.760	2.760	2.763	2.720	2.770	2.755
Zn	CdTe	2.673	2.688	2.676	2.660	2.665	2.671	2.658	2.674
Hg	CdTe	2.800	2.801	2.800	2.800	2.799	2.801	2.790	
Cd	HgTe	2.804	2.803	2.804	2.805	2.805	2.804	2.813	
Hg	ZnTe	2.750	2.735	2.754	2.753	2.757	2.715	2.758	2.748
Zn	HgTe	2.671	2.685	2.674	2.659	2.664	2.671	2.662	2.673
As	AlP	2.425	2.418	2.429	2.427	2.428	2.406	2.427	2.422
P	AlAs	2.392	2.399	2.387	2.387	2.389	1.394	2.390	2.395
As	GaP	2.417	2.409	2.424	2.417	2.420	2.396	2.416	2.414
P	GaAs	2.386	2.393	2.380	2.382	2.383	2.389	2.386	2.387
As	InP	2.596	2.589	2.599	2.598	2.600	2.579	2.598	2.595
P	InAs	2.561	2.568	2.558	2.557	2.558	2.563	2.560	2.562
Sb	AlAs	2.584	2.566	2.597	2.587	2.592	2.539	2.577	2.574
As	AlSb	2.506	2.522	2.496	2.495	2.498	2.511	2.514	2.510
Sb	GaAs	2.569	2.553	2.584	2.571	2.576	2.524	2.546	2.564
As	GaSb	2.501	2.516	2.489	2.492	2.495	2.508	2.525	2.505
Sb	InAs	2.747	2.730	2.754	2.750	2.754	2.705	2.733	2.739
As	InSb	2.669	2.683	2.663	2.658	2.662	2.672	2.683	2.667
Sb	AlP	2.555	2.529	2.569	2.561	2.568	2.488	2.550	2.542
P	AlSb	2.440	2.462	2.426	2.425	2.430	2.447	2.448	2.444
Sb	GaP	2.526	2.503	2.551	2.529	2.537	2.461	2.504	2.519
P	GaSb	2.431	2.451	2.414	2.418	2.422	2.440	2.454	2.436
Sb	InP	2.712	2.687	2.720	2.719	2.725	2.654	2.702	2.700
P	InSb	2.599	2.619	2.591	2.585	2.590	2.604	2.611	2.597
Se	ZnS	2.420	2.409	2.421	2.424	2.426	2.396	2.426	2.420
S	ZnSe	2.367	2.376	2.365	2.360	2.364	2.370	2.363	2.367
Te	ZnSe	2.586	2.569	2.588	2.589	2.592	2.540	2.591	2.584
Se	ZnTe	2.501	2.517	2.497	2.490	2.494	2.504	2.495	2.502
Te	ZnS	2.543	2.513	2.544	2.552	2.558	2.478	2.557	2.539
S	ZnTe	2.406	2.429	2.400	2.390	2.396	2.410	2.396	2.407

TABLE IV. Deviations of the calculated impurity bond lengths (in Å) as compared with the experimental values from EXAFS.

Impurity	Host	A	C	D1	FP[a]	MZ	SSHS[a]	Expt.
In	GaAs	−0.028	−0.014	−0.025	−0.017	−0.031	−0.009	2.587[b]
Ga	InAs	0.004	−0.001	−0.005	−0.006	0.008	0.005	2.487[b]
Cd	ZnTe	0.004	0.008	−0.008	0.018	0.003	0.012	2.752[c]
Zn	CdTe	−0.01 to −0.02	0.0 to −0.01	−0.02 to −0.03	−0.02 to −0.03	−0.01 to −0.02	0.0 to −0.01	2.68 to 2.69[c]
Te	ZnSe	−0.009	−0.07	−0.003	−0.004	−0.011	−0.004	2.595[d]
Se	ZnTe	0.005	0.01	−0.006	−0.001	0.006	0.004	2.496[d]
Mean absolute deviation		0.011	0.006	0.012	0.012	0.012	0.006	

[a]SSHS, Ref. 8.
[b]Ref. 1.
[c]Quoted in Ref. 3.
[d]Reference 14.

A by 0.025 Å. Model A produces about the same d_1 values as model D1, where the maximum difference in d_1 is only 0.015 Å. Martins and Zunger[3] used the same model as model A; however, their analytic expression for δ is different from that given by Eq. (21). Nevertheless, numerical results indicate that these two calculations agree to 0.01 Å. The slightly different forms of strain energies used in models D2 and D1 only introduce a small change in d_1, with the largest difference being less than 0.01 Å. The first-shell continuum model (model E) allows too little relaxation, so while the other models produce a ratio δ/δ_0 ranging from 0.6 to 0.8, model E only ranges from 0.4 to 0.6. The reason that the fixed boundary in model A works is that the effective shear coefficient C (see Table I) characterizing the strain energy in the elastic continuum is large. However, model B is too rigid and does not provide enough buffer between the impurity bond and the fixed boundary.

The comparison of the theoretical results with the available experimental data in Table IV indicates that models B and E are the least accurate. Models A, D1, that of Martins and Zunger (MZ), and the full theory are comparable in that all have an average absolute deviation of 0.012 Å, which is about the experimental uncertainty in extended x-ray-absorption fine structure (EXAFS). The agreement between theory and experiment, however, is not uniform. The most surprising result in Table IV is that the simple spring model (model C) and its cruder version used by Shih *et al.*[8] ($\alpha=\alpha_l$ so $\delta/\delta_0=0.75$, labeled as SSHS) have the smallest variance in d_1, about 0.006 Å. We know this does not imply that the simple spring model represents the real picture of bond-length relaxation. For example, if we let all the shear coefficients be 0, i.e., $\beta=C=0$ in our model, then as the range of the boundary is gradually extended, the local bond length will eventually relax to the impurity bond length $d_1=d_I$, or $\delta=\delta_0$. This can be seen in model A from Eq. (21), where δ reduces to $\delta_0/(1+\alpha/6\alpha_l)$, rather than $\delta_0/(1+\alpha/3\alpha_l)$ as predicted by model C, and in model D from Eqs. (27) and (30), δ becomes δ_0, if the continuum is taken to be shearless. Considering that various effects are included that may mask the absolute accuracy of d_1 predictions (e.g., while low-temperature bond lengths are used in the calculation, the room-temperature values of elastic constants are adopted), the agreement of various models with experiments in Table IV should be regarded as excellent. There are, however, many other impurity systems in which the simple-spring-model predictions differ considerably from other models, as is shown in Fig. 3, where δ is plotted against δ_0 for the full theory. Those points that deviate significantly from the 0.75-slope line are the systems with (As,Sb) or (P,Sb) substitutions. Additional EXAFS measurements on these systems are needed to test these predictions.

C. Mixing enthalpies

Table V lists the mixing enthalpy parameters Ω (in kcal/mole) for a number of alloys estimated from Eq. (41) for all the models considered, along with other theoretical[3,16–18] and experimental values.[16,19] As already dis-

TABLE V. Mixing enthalpy parameter Ω (in kcal/mole) estimated from the full perturbation theory and several valence force models discussed in Sec. III, and comparison with experiments and other theories.

	A	B	C	D1	D2	E	FPT	MZ[a]	DL[b]	FM[c]	VV[d]	Expt.[e]
(Ga,Al)P	0.00	0.01	0.00	0.01	0.01	0.01	−0.05					
(Ga,Al)As	0.00	0.00	0.00	0.00	0.00	0.00	−0.07	0.02	0.02	0.03	0.11	0.0
(Ga,Al)Sb	0.02	0.03	0.02	0.02	0.02	0.03	−0.15	0.02	0.02	0.03		0.0
(Ga,In)P	3.76	4.79	3.0	3.29	3.39	5.24	2.54	4.56	3.63	2.94		3.25,3.5
(Ga,In)As	2.97	3.76	2.36	2.61	2.69	4.14	1.60	2.49	2.81	2.42	1.25	1.65, 2.0, 3.0
(Ga,In)Sb	2.22	2.83	1.77	1.95	2.00	3.09	0.81	2.53	1.85	1.83		1.47, 1.9
(In,Al)P	3.24	4.22	2.77	2.78	2.87	4.60	2.55					
(In,Al)As	2.86	3.65	2.32	2.49	2.56	3.93	2.17	3.60	2.81	2.37		2.5
(In,Al)Sb	1.81	2.33	1.49	1.57	1.61	2.50	1.36	2.06	1.46	1.45		0.6
(Cd,Zn)Te	1.80	2.43	1.73	1.43	1.49	2.45	1.24	2.12	1.97	1.63		1.34
(Hg,Cd)Te	0.00	0.00	0.00	0.00	0.00	0.00	−0.07					0.7, 1.4
(Hg,Zn)Te	1.63	2.20	1.56	1.30	1.36	2.23	1.50	1.91	1.81	1.48		3.0
Al(P,As)	0.81	1.03	0.65	0.71	0.73	1.14	0.76					
Ga(P,As)	0.95	1.18	0.70	0.86	0.87	1.32	0.94	1.15	0.98	0.66	0.12	0.4, 1.0
In(P,As)	0.60	0.78	0.52	0.51	0.53	0.84	0.57	0.72	0.58	0.52		0.4
Al(As,Sb)	4.31	5.45	3.38	3.80	3.88	5.92	4.09					
Ga(As,Sb)	3.77	4.69	2.81	3.40	3.46	5.22	3.67	4.58	3.35	2.76		4.0, 4.5
In(As,Sb)	2.61	3.39	2.23	2.24	2.31	3.67	2.52	2.89	2.29	2.17	6.65	2.25, 2.9
Al(P,Sb)	8.60	10.99	6.99	7.54	7.73	12.00	8.32					
Ga(P,Sb)	8.54	10.61	6.36	7.72	7.88	11.66	8.66					
In(P,Sb)	5.87	7.64	5.08	4.99	5.15	8.04	5.76					
Zn(S,Se)	1.04	1.39	0.98	0.85	0.90	1.49	0.90					
Zn(S,Se)	1.04	1.39	0.98	0.85	0.90	1.49	0.90					
Zn(Se,Te)	2.47	3.27	2.23	2.09	2.16	3.63	2.26	2.91	3.11	2.12	3.12	1.55
Zn(S,Te)	7.02	9.34	6.45	5.80	6.02	9.72	6.20					

[a]Reference 3, column A of Table III.
[b]Reference 16.
[c]Reference 18.
[d]Reference 17.
[e]References 16 and 19.

cussed, the chemical terms reduce the excess energies in the cation impurities and increase them for anion impurities. The corresponding changes in Ω are the differences between the FPT and model D2. We note that the reductions of Ω for the (Ga,In) alloys are very large (> 1 kcal/mole) and also significant for (In,Al) alloys. However, the increases in Ω for the anion substitutional alloys are not as large. Also, the Ω differences between models D1 and D2 are less than 10%. Model A produces Ω values about 20% larger than model D1, model B in turn is 20% higher than model A, and model E is 10% higher than model B. The Ω values in the simple spring model (model C) are seen to be about the same as model D1, although the differences among systems can be positive or negative. Although MZ used the same strain model as model A, their Ω values do not agree with our model A values because their way of estimating Ω is different. In fact, the values of MZ are closer to model B than to A.

To distinguish the quantitative nature of different theoretical models, we note that there are also important factors that may mask the comparison between theory and experiment for Ω. One important factor is that the mixing enthalpies extracted from phase-diagram analysis are sensitive to sample and experimental conditions. These Δ_{H_m} contain contributions from various nonideal structures such as vacancies, impurities, dislocations, grain boundaries, and surface conditions, in addition to the ideal ΔH_m considered here for solid solutions. Thus, our theoretical ΔH_m should represent a lower bound. Another complication comes from the version of the theory of solid solution adopted. The theory used for analysis so far assumes a regular solid solution with second-neighbor pair interactions as was outlined in Sec. V. Recent experimental[20–22] and theoretical[23] studies

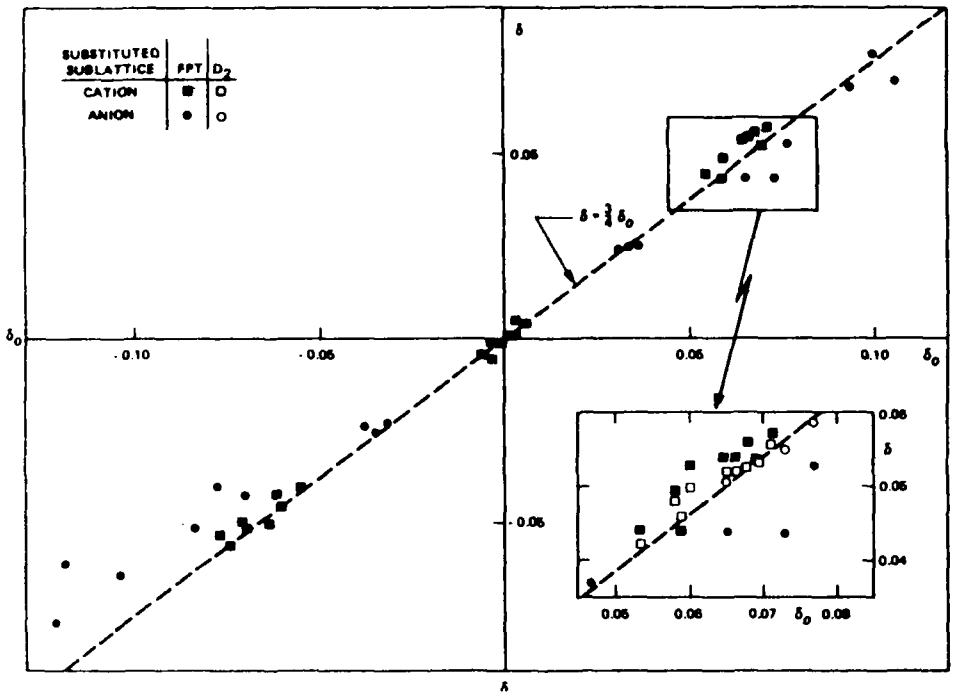


FIG. 3. Calculated bond-relaxation parameter δ from FPT and D2 as a function of δ_0. The $\delta=\frac{3}{4}\delta_0$ curve corresponds to the theory of Shih, Spicer, Harrison, and Sher (Ref. 8).

have suggested the possibility of compositional clustering or long-range correlations in alloys. Extending the theory to include such effects will alter the simple results for ΔH_m from Eq. (41). Moreover, there is evidence from the composition variation of the alloy hardness[24] and from the optical-phonon frequencies[25] that the shear coefficients of alloys increase near the center of the composition range. This will cause the effective continuum shear coef-

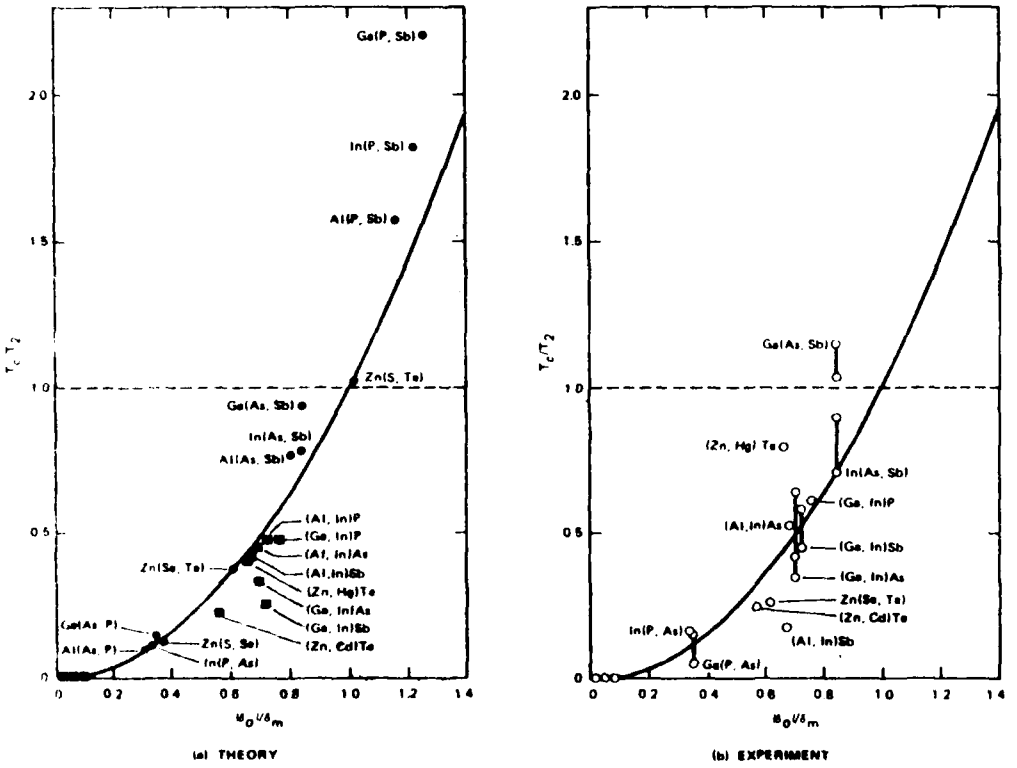


FIG. 4. Plot of T_c/T_2 as a function δ_0/δ_m for T_c obtained from (a) FPT and (b) the experimental Ω in Table V. In (a), the circles are for anion alloys, the squares for cation substitution. The solid lines in both parts correspond to the simple theory discussed in Appendix C. The dashed lines at $T_c/T_2=0$ separate the miscible from immiscible groups.

TABLE VI. Comparison of the critical temperature T_c of mixing and melting temperatures of the constituents T_1 and T_2, in the order of their appearance in the parentheses. Also tabulated are the averaged absolute values of $|\delta_0|$ and the ratio $|\delta_0|/\delta_m$.

System	$\lvert\delta_0\rvert$ (%)	T_c (K)	T_c/T_1	T_c/T_2	$\lvert\delta_0\rvert/\delta_m$
(Al,Ga)Sb	0.6	0	0	0	0.067
(Al,Ga)As	0.1	0	0	0	0.009
(Al,Ga)P	0.3	0	0	0	0.029
(Al,In)Sb	5.5	342	0.25	0.42	0.679
(Ga,In)Sb	5.8	204	0.21	0.25	0.716
(Al,In)As	6.8	547	0.29	0.45	0.687
(Ga,In)As	6.9	403	0.23	0.34	0.697
(Al,In)P	7.1	642	0.36	0.48	0.732
(Ga,In)P	7.4	639	0.43	0.48	0.763
In(P,As)	3.2	144	0.11	0.12	0.330
Ga(P,As)	3.7	236	0.14	0.15	0.352
Ga(P,As)	3.5	191	0.10	0.11	0.307
In(As,Sb)	6.8	635	0.52	0.79	0.840
Ga(As,Sb)	7.6	924	0.53	0.94	0.844
Al(As,Sb)	8.1	1030	0.56	0.78	0.810
In(P,Sb)	9.9	1450	1.08	1.82	1.222
Ga(P,Sb)	11.3	2180	1.25	2.21	1.256
Al(P,Sb)	11.6	2095	1.19	1.58	1.116
(Cd,Hg)Te	0.3	0	0	0	0.033
(Zn,Hg)Te	6.0	377	0.25	0.40	0.659
(Zn,Cd)Te	6.2	312	0.21	0.23	0.564
Zn(S,Se)	4.7	226	0.11	0.13	0.362
Zn(Se,Te)	7.2	569	0.32	0.38	0.615
Zn(S,Te)	11.9	1560	0.74	1.03	1.017

ficient C in Eq. (14) to be composition dependent, which will cause Ω to increase. Despite these uncertainties, useful comparisons across the board in Table V can still be made.

Based on the above considerations, we can conclude that models B, E, and MZ predict Ω values that are too high. We should emphasize that all the Ω numbers for models from A through MZ are directly calculated without any adjustable parameters. The fact that models A, C, and D1,D2, and the FPT agree with the experiment as well as or even slightly better than the one-parameter theories, the delta-lattice (DL) model[16] and the model of Fedder and Muller[18] (FM), is already quite an accomplishment. The few numbers taken from Van Vechten's calculations[17] (VV) indicate that the dielectric model predicts results at larger variance with experiments. There are two important implications about the FPT that can be drawn from Table III. First, the theory predicts a small but negative Ω value for several alloys. This not only means that there is no miscibility gap in these alloys but also implies a tendency toward ordering, in which the substitutional atoms tend to be surrounded by different second-neighbor species. For stoichiometric compositions, this implies a tendency toward compound formation. Secondly, the FPT tends to predict smaller Ω values than observed experimentally, which should be expected according to our discussion. To the extent that the FPT predicts the correct ΔH_m values for an ideal solution, the difference $\Delta H_m^{\text{expt}} - \Delta H_m$ may be attributed to imperfect conditions and deviations from the ideal solution theory.

Finally, the calculated Ω values in Table V provide some guidance in separating the completely miscible alloys from immiscible ones.[26,27] In a true random alloy, the criteria[28] for alloy mixing at a temperature T is that $T \geq T_c$, where the critical temperature T_c is given by $\Omega/(2R_g)$, with R_g being the universal gas constant.[29] For an $A_xB_{1-x}C$ alloy to be miscible throughout the whole concentration range, the requirement is that both the melting temperatures T_1 and T_2 of the pure AC and BC compounds be greater than T_c. Table VI lists the values of T_c associated with the Ω values in the FPT, the ratios T_c/T_1 and T_c/T_2, and the average absolute values of δ_0 for the alloys considered in Table V. In Table VI, T_c is set equal to zero if Ω is negative and T_2 is chosen to be the lower value of the two melting temperatures, so the criterion for not having a miscibility gap is $T_c/T_2 < 1$.

There is an empirical rule[26] stating that this will be satisfied if the lattice mismatch $|\delta_0|$ between the two alloy components is less than 7.5%. However, we find that (see Appendix C) a more precise rule is $|\delta_0| > \delta_m$, where [illegible] and X_m is the ratio of the rms bond-length amplitude fluctuation to the bond length at the melting temperature T_m. The values of T_m for the compounds involved[40] and the associate X_m values estimated from Eqs. (C2) and (C3) are tabulated in Table I. The model used in Appendix C yields [illegible] $(\delta_0/\delta_m)^2$. This suggests that [illegible] to plot T_c/T_2 as a function of $|\delta_0|/\delta_m$, [illegible] in Fig. 4 for the T_c calculated from FPT. [illegible] similar to the ΔE versus δ_0 curve in Fig. 2 because [illegible] Ω is proportional to the sum of the ΔE [illegible] constituents [see Eq. (41)]. However, if [illegible] against δ_0 alone, the FPT points are [illegible] scattered, and those of SSHS would not even [illegible] smooth simple quadratic form because the lower melting temperature T_2 is not a smooth function of $|\delta_0|$. [illegible] suggests that $\delta_0/\delta_m < 1$ is a better criterion [illegible] 0.075. Figure 3 also clearly shows the chem- [illegible] all the cation-substitution alloy points lie [illegible] curve and all the anion-substitution alloys [illegible] on or above the curve, corresponding [illegible] shifts in ΔE due to the chemical [illegible] the curve based on the SSHS model is an [illegible] representation. From the figure, we [illegible] should have miscibility gaps and [illegible] alloys are predicted to be miscible, although on the borderline, *because the actual mixing enthalpies are* larger than these ideal calculated values. The figure also shows that (Zn,S,Te) has a miscibility gap but a smaller [illegible] the (P,Sb) alloys, despite the fact [illegible] is larger. All these predictions are [illegible] with the available experimental evidence.

VII. SUMMARY AND CONCLUSION

[illegible] theory of defect substitution en- [illegible]. The substitution energy is compactly [illegible] energy Δ_E and a distortion [illegible] [see Eq. (2)]. However, a [illegible] requires an improve- [illegible] aspects of Harrison's bonding [illegible] the [illegible] constants.[4] The most in- [illegible] this theory presented in this paper [illegible] which enables us to absorb the [illegible] coefficients into the calculation and, [illegible] the chemical effects. The ori- [illegible] on impurity bond relaxation [illegible] mechanisms [see Eq. (17)]: a [illegible] either helps or hinders lattice re- [illegible] whether it has the same or oppo- [illegible] bond-length *difference* $d-d_I$ between [illegible] chemical energy that depends on [illegible] between the impurity and [illegible] effective elastic force constant H [illegible] tends to restrain the lattice from [illegible] the effect of boundaries between the [illegible] impurity and the rest of the elastic [illegible] based on the valence force field[10] are derived and their results are compared with the full perturbation theory and available experimental data. We found at least five models, including the FPT, that produce the correct impurity bond lengths with variances for the compounds studied about equal to the experimental uncertainties in EXAFS[1,4,14] (~0.01 Å). However, some models are oversimplified and will certainly not predict other properties equally well. However, more experimental lattice constant measurements to further test the theory, particularly on (As,Sb) and (P,Sb) substitution systems for which there are larger differences between different models, are needed. It would also be instructive to see if the predicted reversal for (Hg,Cd)Te is found.

The excess energies of impurity substitution are also shown to provide good estimates of the mixing enthalpies Ω of pseudobinary alloys. The chemical shifts are found to have a negative net contribution to Ω for most cation substitutions, but positive contributions for anion substitutions. The chemical reduction of Ω in (Ga,In) alloys is larger than 1 kcal/mole (30–100 %). Several VFF models and the full perturbation theory produce results for Ω that are as good as the best theories with one adjustable parameter. However, the full theory tends to yield answers on the low side of the experimental values, which we argue is as it should be because there are nonideal structures that also contribute to Ω. The calculated Ω values and the melting temperatures are used to predict the existence of alloy miscibility gaps, and the results correlate well with experiments.

Finally, we wish to comment on the accuracy of the theories that are connected to the present model. The perturbation theory has already been stretched beyond its expected region of validity and predicts d_1 to within experimental uncertainties (~0.01 Å) even for cases with large bond-length differences ($\delta_0 \sim 0.1$). The accuracy can only be improved if the full nonperturbation theory outlined in Sec. II is used. This calculation is needed for the strong substitution cases that were not considered in this paper: examples are (B,Ga), (B,In), (N,P), (N,As), (N,Sb), (O,S), (O,Se), and (O,Te) substitutions. Although we believe that for the properties treated, the model with a continuum attached to the second shell is as accurate as the perturbation theory used, it remains to be seen if this is true for other properties, especially strain coefficients. Finally, the present theory has been extended to study alloys[5,6] by embedding clusters in an effective medium. This enables us to study the bond length and energy variations throughout the whole concentration range. However, a quantitative calculation still awaits an improvement of the accuracy of Harrison's theory. A similar procedure is also being extended to a study of the alloy electronic structure for which a cluster CPA (coherent-potential approximation) involving both potential and structural disorder[11] will be used.

ACKNOWLEDGMENTS

This work benefited from useful discussions with Professor W. A. Harrison and Professor A. Zunger. One of us (A.-B. C.) would like to thank Professor W. E. Spicer for his hospitality. The work is supported in part by U.S.

Air Force Office of Scientific Research Contract No. AFOSR-84-0284 and U.S. Defense Advanced Projects Agency DARPA Contract No. MDA-903-83-C-0108.

APPENDIX A: ELASTIC ENERGY IN CONTINUUM

In Sec. III the elastic energy outside a sphere of radius R centered at the impurity is assumed to be a continuum with a radial displacement $\mathbf{u} \propto \hat{\mathbf{r}}/r^2$. If the displacement at R is u_0, then $\mathbf{u}(\mathbf{r}) = u_0(R^2/r^2)\hat{\mathbf{r}}$. The energy density in the continuum is given by

$$\delta\epsilon(\mathbf{r}) = \tfrac{1}{2}C_{11}(e_{xx}^2+e_{yy}^2+e_{zz}^2) + C_{12}(e_{yy}e_{zz}+e_{xx}e_{yy}+e_{xx}e_{zz}) + \tfrac{1}{2}C_{44}(e_{xy}^2+e_{yz}^2+e_{zx}^2) ,$$

where

$$e_{xx} = \frac{\partial u_x}{\partial x} = R^2u_0(r^2-3x^2)/r^5 ,$$

$$e_{xy} = \frac{\partial u_x}{\partial_x} + \frac{\partial u_y}{\partial x} = -6R^2u_0xy/r^5, \ldots .$$

Thus, the total elastic energy in the continuum is

$$\begin{aligned}\Delta_{\rm dis}^{(\rm out)} &= \int_R^\infty \delta\epsilon(\mathbf{r})d^3r \\ &= 4\pi R u_0^2(\tfrac{2}{5}C_{11}-\tfrac{2}{5}C_{12}+\tfrac{6}{5}C_{44}) \\ &= CRu_0^2 ,\end{aligned}$$

where the effective shear coefficient is given by

$$C = \pi(1.6C_{11}-1.6C_{12}+4.8C_{44}) .$$

APPENDIX B: DISTORTION ENERGY

In this Appendix we count the detailed contributions of the bond-stretching terms $\Delta(\mathbf{r}_i\cdot\mathbf{r}_i)$ and "bond-bending" terms $\Delta(\mathbf{r}_i\cdot\mathbf{r}_j)$ for $i\neq j$ in VFF [Eq. (13)] that enter Eq. (16) in FPT and in the VFF models in Sec. III.

1. α and β terms from the first-shell bonds

The four bond vectors pointing away from the central impurity according to Fig. 1 are

$$\mathbf{r}_1 = (1-\delta, 1-\delta, 1-\delta)d/\sqrt{3} ,$$

$$\mathbf{r}_2 = (1-\delta, -1-\delta, -1-\delta)d\sqrt{3}, \ldots .$$

Thus, $\Delta(\mathbf{r}_1\cdot\mathbf{r}_1) = -2\delta d^2$ and $\Delta(\mathbf{r}_1\cdot\mathbf{r}_2) = -\frac{2}{3}\delta d^2$. The α terms contribute

$$4\times 3\alpha_1(-2\delta d^2)^2/8d^2 = 6\alpha_1\delta^2d^2 ,$$

and the β terms contribute

$$6\times 3\beta_1(-\tfrac{2}{3}\delta d)^2/8d^2 = \beta_1\delta^2d^2 .$$

If an A atom is replaced by a B atom, as was done in FPT, the α_1 and β_1 are replaced by α and β, respectively.

2. α terms from the second-shell bonds, β terms between the first- and second-shell bonds and among the second-shell bonds

For these terms we need to consider the four bond vectors pointing away from C. They are

$$\mathbf{r}_1 = (-1+\delta, -1+\delta, -1+\delta)d/\sqrt{3} ,$$

$$\mathbf{r}_2 = (-1+\delta, 1+\delta+\gamma, 1+\delta+\gamma)d/\sqrt{3} ,$$

$$\mathbf{r}_3 = (1+\delta+\gamma, -1+\delta, 1+\gamma+\delta)d/\sqrt{3}, \ldots .$$

Then $\Delta(\mathbf{r}_2\cdot\mathbf{r}_2) = \frac{2}{3}(\delta+2\gamma)d^2$, $\Delta(\mathbf{r}_1\cdot\mathbf{r}_2) = -\frac{2}{3}(\delta+\gamma)d^2$, and $\Delta(\mathbf{r}_2\cdot\mathbf{r}_3) = \frac{2}{3}\delta d^2$. Thus, the α terms from the second-shell bonds become

$$4\times 3\times 3\alpha[\tfrac{2}{3}(\delta+2\gamma)d^2]^2/8d^2 = 2\alpha(\delta+2\delta)^2d^2 ,$$

the β term between the first- and second-shell bonds are

$$4\times 3\times 3\beta[\tfrac{2}{3}(\delta+\gamma)d^2]^2/8d^2 = 2(\gamma+\delta)^2\beta d^2 ,$$

and the β terms among the second-shell bonds are

$$4\times 3\times 3\beta(\tfrac{2}{3}\delta d^2)^2/8d^2 = 2\delta^2\beta d^2 .$$

3. α terms for the third-shell bonds, β terms between the second- and third-shell bonds and among the third-shell bonds adjacent to the second-shell atoms

For these terms we need to consider the bond vectors pointing away from B in Fig. 1. They are

$$\mathbf{r}_2 = (1-\delta, -1-\gamma-\delta, -1-\gamma-\delta)d/\sqrt{3} ,$$

$$\mathbf{r}_1 = (1+\gamma', 1+3\gamma'-\gamma, 1+3\gamma'-\gamma)d/\sqrt{3} ,$$

$$\mathbf{r}_3 = (-1-\gamma'', 1+3\gamma''-\gamma, -1+\gamma''-\gamma)d/\sqrt{3} ,$$

and

$$\mathbf{r}_4 = (-1-\gamma'', -1+\gamma''-\gamma, 1+3\gamma''-\gamma)d\sqrt{3} .$$

Thus, we have

$$\Delta(\mathbf{r}_1\cdot\mathbf{r}_2) = -\frac{d^2}{3}(3\delta+5\gamma') ,$$

$$\Delta(\mathbf{r}_2\cdot\mathbf{r}_3) = \frac{d^2}{3}(\delta+2\gamma-5\gamma'') ,$$

$$\Delta(\mathbf{r}_3\cdot\mathbf{r}_4) = -\tfrac{2}{3}d^2\gamma'' ,$$

$$\Delta(\mathbf{r}_1\cdot\mathbf{r}_3) = \frac{d^2}{3}(-2\gamma-\gamma'+3\gamma'') ,$$

$$\Delta(\mathbf{r}_3\cdot\mathbf{r}_3) = \frac{d^2}{3}(6\gamma'') ,$$

and

$$\Delta(\mathbf{r}_1\cdot\mathbf{r}_1) = \frac{d^2}{3}(14\gamma'-4\gamma) .$$

For model A, $\gamma' = r'' = 0$, so the α terms of the third-shell bonds become

$$4\times 3\times 3\alpha[2(\Delta\mathbf{r}_3\cdot\mathbf{r}_3)^2+(\Delta\mathbf{r}_1\cdot\mathbf{r}_1)^2]/8d^2 = 8\alpha\gamma^2d^2 ,$$

the β terms between the second- and third-shell bonds are

$$4\times3\times3\beta[(\Delta\mathbf{r}_1\cdot\mathbf{r}_2)^2+2(\Delta\mathbf{r}_2\cdot\mathbf{r}_3)^2]/8d^2$$
$$=\tfrac{9}{2}\beta d^2[\delta^2+\tfrac{2}{9}(\delta+2\gamma)^2]=\tfrac{9}{2}\beta\delta^2d^2+\beta(\delta+2\gamma)^2d^2\ ,$$

and the β terms among the third-shell bonds adjacent to the second-shell atoms are

$$4\times3\times3\beta[2(\Delta\mathbf{r}_1\cdot\mathbf{r}_3)^2/(8d^2)+(\Delta\mathbf{r}_3\cdot\mathbf{r}_4)^2]=4\beta\gamma^2d^2\ .$$

For continuum and the only contribution from this group are the β terms between the second- and third-shell bonds. Since the displacements in the continuum are proportional to $1/R^2$, $\gamma'=8\sqrt{2}\gamma/(19\sqrt{19})$ and $\gamma''=8\sqrt{2}\gamma/(11\sqrt{11})$. Thus, these β terms become

$$\frac{9}{2d^2}\beta\left[\frac{d^4}{9}(3\delta+5\gamma')^2+\frac{2d^4}{9}(\delta+2\gamma-5\gamma'')^2\right]$$
$$=\tfrac{1}{2}\beta\left[\left|3\delta+\frac{40\sqrt{2}}{19\sqrt{19}}\gamma\right|^2+2\left|\delta+2\gamma-\frac{40\sqrt{2}}{11\sqrt{11}}\gamma\right|^2\right]d^2\ .$$

4. β terms for bonds adjacent to the third-shell atoms

These terms only enter model A, so $r'=r''=0$. There are two different groups, one like those adjacent to C' and another like those meeting at C''. The four bond vectors pointing away from C' are

$$\mathbf{r}_1=(-1,-1+\gamma,-1+\gamma)d/\sqrt{3}\ ,$$
$$\mathbf{r}_2=(-1,1,1)d/\sqrt{3}\ ,$$
$$\mathbf{r}_3=(1,-1,1)d/\sqrt{3}\ ,$$

and

$$\mathbf{r}_4=(1,1,-1)d/\sqrt{3}\ .$$

Thus, the only contribution from this group is

$$4\times3\times3\beta[\Delta(\mathbf{r}_1\cdot\mathbf{r}_2)]^2/8d^2=2\beta\gamma^2d^2\ .$$

The four bond vectors around C''' are

$$\mathbf{r}_1=(-1+\gamma,-1-\gamma,-1)d/\sqrt{3}\ ,$$
$$\mathbf{r}_2=(-1,1,1)d/\sqrt{3}\ ,$$
$$\mathbf{r}_3=(1,-1+\gamma,1+\gamma)d/\sqrt{3}\ ,$$

and

$$\mathbf{r}_4=(1,1,-1)d/\sqrt{3}\ ,$$

which only results in the first-order term $\Delta(\mathbf{r}_2\cdot\mathbf{r}_3)=\frac{2}{3}\gamma d^2$. Thus, the group contributes to

$$4\times3\times3\beta[2(\Delta\mathbf{r}_2\cdot\mathbf{r}_3)^2]/8d^2=4\beta\gamma^2d^2\ ,$$

and the combined contribution from these two groups is $6\beta\gamma^2d^2$.

APPENDIX C: CRITERION OF MISCIBILITY

Starting with Eq. (24) and using the SSHS model $\alpha=\alpha_1$, one finds the mixing enthalpy parameter Ω to be

$$\Omega=\tfrac{3}{2}\bar{\alpha}(d_{AC}-d_{BC})^2N_0\ , \tag{C1}$$

where N_0 is Avagodro's number and $\bar{\alpha}=\frac{1}{2}(\alpha_{AC}+\alpha_{BC})$. Then relate the mean-square bond-length fluctuation $\langle\zeta^2\rangle$ at the melting temperature T_m to T_m for a compound by equating the average potential energy per unit cell to half of the thermal energy:

$$\langle V_{\rm pot}\rangle\simeq4\times\tfrac{3}{2}\alpha\langle\zeta^2\rangle=\tfrac{1}{2}(2\times3k_BT_m)\ , \tag{C2}$$

where k_B is the Boltzmann constant. Defining a Lieder-mann ratio of melting χ_m by

$$(\langle\zeta^2\rangle)^{1/2}\approx\chi_md \tag{C3}$$

and choosing the mixing criterion to be $T_c/T_m<1$, where T_m now is the smaller value of the two melting temperatures of the constituent compounds, we require that

$$\frac{T_c}{T_m}=\frac{\Omega k}{4R_g\alpha\chi_m^2d^2}=\frac{3}{8}\frac{(d_{AC}-d_{BC})^2}{\chi_m^2d^2}<1 \tag{C4}$$

or

$$|\delta_0|/\delta_m<1\ , \tag{C5}$$

where $\delta_m=1.63\chi_m$ and $|\delta_0|$ is the percentage bond-length difference.

[1] J. C. Mikkelsen and J. B. Boyce, Phys. Rev. Lett. **49**, 1412 (1982); Phys. Rev. B **28**, 7130 (1983).

[2] A. Zunger and J. E. Jaffe, Phys. Rev. Lett. **51**, 662 (1983).

[3] J. L. Martins and A. Zunger, Phys. Rev. B **30**, 6217 (1984).

[4] A. Balzarotti, A. Kisiel, N. Motta, M. Zimmal-Starnawska, M. T. Czyzyk, and M. Podgorny, Phys. Rev. B **30**, 2295 (1984).

[5] A.-B. Chen and A. Sher, Microscience **3**, 1 (1984).

[6] A. Sher, A. B. Chen, and W. E. Spicer, in *13th International Conference on Defects in Semiconductors*, edited by L. C. Kimerling and J. M. Parsey, Jr. (The Metallurgical Society of AIME, New York, 1985), p. 335.

[7] W. A. Harrison, *Electronic Structure and Properties of Solids* (Freeman, San Francisco, 1980); Microscience **3**, 35 (1983); Phys. Rev. B **27**, 3592 (1983).

[8] K. Shih, W. E. Spicer, W. A. Harrison, and A. Sher, Phys. Rev. B **31**, 1139 (1985).

[9] A. Baldereschi and J. J. Hopfield, Phys. Rev. Lett. **28**, 171 (1972).

[10] R. M. Martin, Phys. Rev. B **1**, 4005 (1970).

[11] The experimental elastic constants were taken from the tabulated values in Refs. 10 and 25 except AlP and AlAs, for which we used the estimated values by J. D. Wiley, in *Semi-*

conductors and Semimetals, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1975), Vol. 10, p. 134.
[12]R. Kubuchi, Physica **103B**, 41 (1981).
[13]A.-B. Chen and A. Sher (unpublished).
[14]J. B. Boyce and J. C. Mikkelsen, unpublished data for Zn(Se,Te). We thank them for their communication prior to publication.
[15]Extrapolated results quoted in Ref. 3.
[16]G. B. Stringfellow, J. Cryst. Growth **27**, 21 (1974); **58**, 194 (1982).
[17]J. A. Van Vechten, in *Semiconductor Handbook*, edited by S. P. Keller (North-Holland, Amsterdam, 1980), Vol. 3, p. 1.
[18]P. A. Fedders and M. W. Muller, J. Phys. Chem. Solids **45**, 685 (1984).
[19]M. B. Panish and Ilegems, Prog. Solid State Chem. **7**, 39 (1972); A. Langier, Rev. Phys. Appl. **8**, 2959 (1973).
[20]T. P. Pearshall, R. Carles, and J. C. Portal, Appl. Phys. Lett. **42**, 436 (1983).
[21]K. Kakimoto and T. Katoda, Appl. Phys. Lett. **40**, 826 (1982).
[22]P. Parayanthal and F. H. Pollak, Phys. Rev. Lett. **52**, 1922 (1984).
[23]G. P. Srivastova, J. L. Martins, and A. Zunger, Phys. Rev. B **31**, 2561 (1985).
[24]N. A. Groyunova, A. S. Barshchvevskii, and D. N. Fretiakov, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1968), Vol. 4, Chap. 1.
[25]S. S. Mitra and N. E. Massa, in *Handbook on Semiconductors*, edited by W. Paul (North-Holland, Amsterdam, 1982), Vol. 1, p. 29.
[26]L. M. Foster, J. Electrochem. Soc. Solid State Sci. Technol. **121**, 1662 (1974).
[27]L. M. Foster and J. F. Woods, J. Electrochem. Soc. Solid State Sci. Technol. **118**, 1175 (1971).
[28]R. Kubo, *Statistical Mechanics* (University of Tokyo Press, Tokyo, 1971), p. 336.
[29]The critical temperature will be reduced by a few percent if more elaborate models are used. See W. Christian, *in Transformations in Metals and Alloys*, 2nd ed., edited by D. W. Hopkins (Pergamon, New York, 1981), p. 194.
[30]*American Institute of Physics Handbook*, 3rd ed., edited by D. E. Gray (McGraw-Hill, New York, 1972), pp. 9—16.
[31]K. C. Hass, R. J. Lempert, and H. Ehrenreich, Phys. Rev. Lett. **52**, 77 (1984).

Reprinted from: Physical Review B, Vol. 33, No. 2, 15 January 1986

Band structures of Si_xGe_{1-x} alloys

Srinivasan Krishnamurthy and A. Sher

A.-B. Chen

PHYSICAL REVIEW B VOLUME 33, NUMBER 2 15 JANUARY 1986

Band structures of Si_xGe_{1-x} alloys

Srinivasan Krishnamurthy and A. Sher
SRI International, Menlo Park, California 94025

A.-B. Chen
Physics Department, Auburn University, Auburn, Alabama 36849
(Received 16 July 1985)

By starting from realistic band structures of the constituent materials, the electronic structure of Si_xGe_{1-x} alloys are obtained in the coherent-potential approximation (CPA). Various quantities, including the bowing parameter of the fundamental band gap and the energies of several optical gaps, the masses, and the linewidths of the E_0 and E_1 transitions, are calculated on the basis of both diagonal and off-diagonal CPA. All of the band-energy and linewidth predictions are in good agreement with experiments. Furthermore, the theory yields an alloy-scattering-limited electron-drift mobility in qualitative agreement with experimental results.

I. INTRODUCTION

Semiconductor alloys offer the freedom to design material properties by choosing appropriate alloy constituents. In some cases, the physical properties of the alloys can be quite different from those of the constituents.[1-4] In recent years, there has been a renewed interest in Si_xGe_{1-x} alloys[5] and superlattices.[5-10] Because silicon is the most technologically advanced semiconductor, the results on Si-Ge systems have many potential applications.

The lattice constants of silicon and germanium differ by $\sim 4\%$. Hence, the strain introduced in the formation of Si_xGe_{1-x} alloys can affect the band structure[10] and the transport properties.[8] Prior authors used virtual-crystal approximation (VCA) (Refs. 11 and 12) and coherent-potential approximation (CPA) (Ref. 13) to study the band structure and related properties. Either because of less accurate band structures of the constituent materials, or because of the approximations involved in the alloy formalism, these calculations predicted only trends of specific quantities, not quantitatively accurate results. Because the s-state site potentials (ϵ_s) for silicon and germanium differ by approximately 1.5 eV, VCA cannot accurately describe effective masses and other finer details of the band structure. Because of the use of poor basis functions, earlier CPA work[13] predicted alloy broadening of conduction-band states substantially differing from experiment. The purpose of this paper is to correct these flaws and treat transport phenomena.

Because of a substantial difference between the site potentials and lattice constants of silicon and germanium, we incorporated both chemical and structural disorder in the calculation of the electronic structure of Si_xGe_{1-x} alloys. Thus, both diagonal and off-diagonal CPA are included in the predicted band structure and related quantities. Parts of the band structure have been used to study the Si $2p$ core exciton[14] and the alloy mobilities.[15] A comprehensive report of the calculations and results is presented here.

The rest of the paper is arranged as follows. The detailed procedure of fitting silicon and germanium band structures is given in Sec. II. The VCA, CPA, and off-diagonal CPA calculations are described in Sec. III. The results and interpretation of the alloy band structures and mobility are given in Sec. IV.

II. BAND-STRUCTURE BASIS

In order to derive an accurate alloy band structure, one must start from a realistic band structure of the constituent materials. Chen and Sher have developed a method,[16] following a prescription of Kane[17] and Chadi,[18] which includes all long-range interactions, and then they have fine-tuned the band structure with an adjustable local Hamiltonian. Because the details have already been published,[16,19] the underlying method will be presented here in brief.

Gaussian orbitals of the type α (α can be s, p_x, p_y, or p_z) for each sublattice in a cell are used to construct the corresponding Bloch basis. In this basis set, the overlap matrix and the Hamiltonian derived from empirical pseudopotentials can be calculated.[17,18] It is possible to cast the problem in a basis set of Gaussian orbitals in which, in crystal units (c.u.), the same exponential factors apply for all III-V compounds.[19] In this universal basis, the overlap matrix and the kinetic energy matrix are the same for all III-V compounds. Then, by a unitary transformation, the basis set is orthonormalized.[20] The Hamiltonian in this new basis set denoted $H_0(\mathbf{k})$. The band structure resulting from this method reproduces the results of elaborate band-structure calculations within a few percent throughout the Brillouin zone (BZ). To establish accurately certain important band-structure features adjacent to the gap, an extra small 8×8 Hamiltonian matrix $H_1(\mathbf{k})$ is added to $H_0(\mathbf{k})$. This $H_1(\mathbf{k})$ has the form of a tight-binding (TB) Hamiltonian, in which only the nearest-neighbor interactions are included, and simulates the effect of nonlocal pseudopotentials and an expanded orbital set. The total Hamiltonian $H(\mathbf{k})$ in this orthonormalized basis set is diagonalized to obtain the band ener-

gies and the corresponding wave functions.

Following this procedure with the same exponential factor $\beta=0.26$ in the Gaussian orbitals for both silicon and germanium, the matrix $H_0(\mathbf{k})$ is obtained. For silicon and germanium, H_1 contains six adjustable parameters: namely, the corrections to the term values Δ_s and Δ_p and to the nearest-neighbor interactions V_{ss}, V_{sp}, V_{xx}, and V_{xy}. The values of Δ_s, Δ_p, V_{ss}, and V_{xx} are determined from fitting the three experimental energy gaps[21-28,39] at $\Gamma(\mathbf{k}=0)$: Γ_2'-Γ_1, Γ_{15}- Γ_{25}', and Γ_2'- Γ_{25}', and the photoelectric threshold values -5.07 and -4.80 eV for silicon and germanium, respectively.[29] The remaining parameters V_{sp} and V_{xy} are obtained from the experimental values[21-28] of the gaps X_{1c}-X_{1v} and L_{1c}-L_{3v}'. Some adjustments in these input quantities are made to obtain an overall good band structure with more accurate effective masses. Table I lists the empirical pseudopotential form factors and the parameters used to obtain the band structure. The calculated band structures and experimental values are given in Table II. From Table II, one can see that an excellent fit to the silicon and germanium band structure is obtained: All the calculated values lie within the experimental uncertainties. The optical difference between L_{1c} and L_{3v}, Γ_{25}' and Γ_{15} are in excellent agreement with the known optical transition values.

Although the calculated effective transverse masses agree very well with experiment, the effective longitudinal mass for germanium is less than the experimental value. This is due mainly to our attempt to have a common β and the choice of local pseudopotentials, causing $H_0(\mathbf{k})$ to be the same in crystal units (c.u.) for both germanium and silicon. Because of the common $H_0(\mathbf{k})$, the alloy disorder is contained in these adjusted parameters. This H_0 would also be useful for the interface and superlattice[31] problems. If we grant ourselves the freedom to adjust V_{xy}, longitudinal effective mass in germanium can be fitted to the experimental value. When V_{xy} is changed, the L_{1c},L_{3v}' will also change. We have chosen not to do this because little is gained for the extra complexity. For an indirect-gap semiconductor, the important effective mass used in transport studies is the conductivity mass

$$3(1/m_{el}^*+2/m_{et}^*)^{-1}.$$

Because $m_{el}^* \gg m_{et}^*$ in germanium, m_c^* will not be much different if a less accurate value of m_{et}^* is used. Moreover, the Si_xGe_{1-x} alloys which have potential device applications are in the silicon-rich region, where the effective mass at the L edge is not expected to affect the further studies.

It is important to note that an excellent fit to the experimental values can be obtained with only seven adjustable parameters $(\beta,\Delta_s,\Delta_p,V_{ss},V_{sp},V_{xx},V_{xy})$, with β being universal in c.u. The calculated band structure of silicon and germanium are shown in Fig. 1(a) and 1(b), respectively. The characteristic indirect gaps are clearly seen.

TABLE I. Pseudopotential form factors and the band parameters (in eV).

Parameter	Silicon	Germanium
$V(\sqrt{3})$	−2.872	−2.872
$V(\sqrt{4})$	0.124	0.124
$V(\sqrt{8})$	0.638	0.638
$V(\sqrt{11})$	0.109	0.109
Δ_s	−16.175	−16.922
Δ_p	−16.109	−14.971
V_{ss}	−0.111	0.131
V_{sp}	0.040	0.150
V_{xx}	0.025	0.030
V_{xy}	0.050	0.100

TABLE II. Calculated symmetry-point band energies of silicon and germanium compared with experiments and empirical-pseudopotential-method (EPM) results. (all energies in eV).

	Silicon		Germanium	
Bands	Calculated	Expt.[a]-EPM[b]	Calculated	Expt.[a]-EPM[b]
Γ_1	−12.60	12.4±0.6	−12.56	−12.6±0.3
L_{2v}	−10.26	−9.3±0.4	−10.74	−10.6±0.5
L_{1v}	−6.99	−6.8±0.2	−7.65	−7.4±0.3
X_{1v}	−8.29	−7.69	−9.20	−8.65
X_{4v}	−2.55	−2.86	−2.55	−3.29
L_{3v}'	−1.11	−1.2±0.2	−1.13	−1.1±0.2
Γ_{25v}'	0.0	0.0	0.0	0.0
L_{1c}	2.24	2.23	0.76	0.76
Γ_{2c}'	4.10	4.00±0.05	0.99	0.99
Γ_{15c}	3.43	3.40	3.24	3.22
X_{1c}	1.34	1.17	0.95	1.16
L_{3c}'	4.34	4.34	4.16	4.25
E_g	1.11	1.11	0.76	0.76
K_0	(0.8,0,0)	(0.8,0,0)	(0.5,0.5,0.5)	(0.5,0.5,0.5)
m_{el}^*	0.89	0.91	1.09	1.59
m_{et}^*	0.16	0.19	0.077	0.082
m_c^*	0.35	0.50	0.28	0.34

[a]References 21–28.
[b]Reference 30.

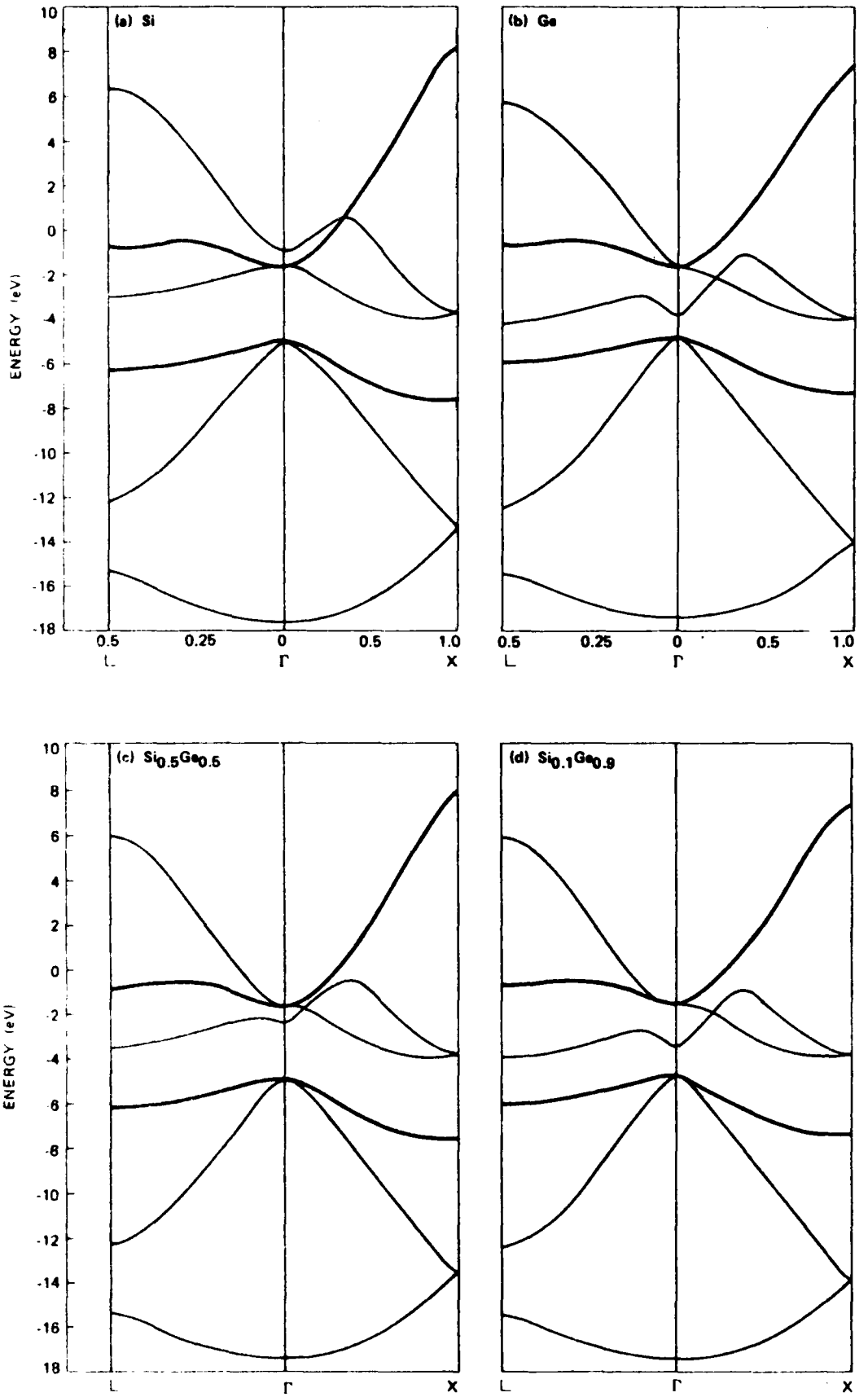


FIG. 1 Calculated VCA band structures of (a) silicon, (b) germanium, (c) $Si_{0.5}Ge_{0.5}$ alloys, and (d) $Si_{0.10}Ge_{0.90}$. (e) (Shown on facing page) shows calculated VCA values (solid) and the experimental values (dashed) of the E_0 and E_0' peak positions plotted as a function of alloy concentration x.

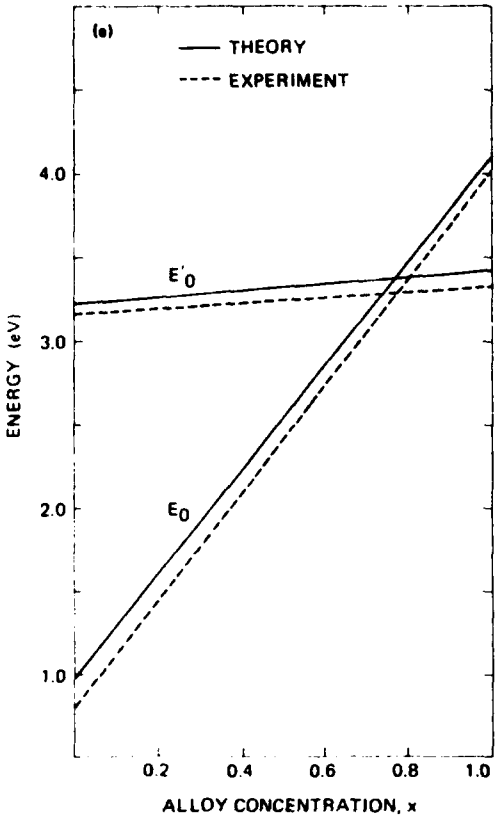


FIG. 1. *(Continued).*

These band structures compare favorably with the best results available, and, in contrast with those obtained in the usual empirical TB approaches, produce good conduction bands. However, if they were included, the spin-orbit interactions (neglected in this work) would add some fine details, especially the splitting between the heavy- and light-hole bands near Γ.

III. ALLOY CALCULATION

A. VCA

Because we have the same $H_0(\mathbf{k})$ matrix for both silicon and germanium, it is only the H_1 matrices of the constituents in scaled VCA which distinguishes them. In this approximation, the diagonal elements of the alloy Hamiltonian $\overline{H}(\mathbf{k})$ are simply the concentration-weighted average of the corresponding elements of the pure silicon and germanium Hamiltonians, whereas the off-diagonal elements of $\overline{H}(\mathbf{k})$ are obtained by assuming a $1/d^2$ dependence. $\overline{H}(\mathbf{k})$ can be diagonalized to obtain the VCA band structure for various concentrations x. The VCA band structures for $x=0.1$ and 0.5 are shown in Fig. 1(c) and 1(d), respectively.

B. Diagonal CPA

An earlier work on CPA band structure of SiGe alloys[11] is based on a local but energy-dependent pseudopotential approximation. While the value of the principal scattering potential parameter, the difference between the *s*-electron site potentials in Si and Ge, was 1.49 eV, close to our value 1.46 eV, the calculation predicted too-large linewidths in the E_0 spectrum and essentially no effect on the electron mobility. With the availability of a set of good basis functions and constituent band structures, more realistic band structures of the alloy can be obtained.

In the current model, we have a TB Hamiltonian, which contains matrix elements to all ranges. The simplest alloy model is to assume that the important disorder resides only in the diagonal matrix elements ϵ_s and ϵ_p. In our model, the ϵ_s^{Si} and ϵ_s^{Ge} differ by 1.46 eV, whereas ϵ_p^{Si} and ϵ_p^{Ge} differ by 0.21 eV. For the present, we neglect the disorder in the off-diagonal element. Mathematically, we have

$$H_{\mathrm{alloy}}=\overline{H}+\sum_l V_l\ , \tag{1}$$

where l is a fcc lattice vector identifying a site, and V_l is the 8×8 diagonal matrix with elements $U_s=\epsilon_s-\bar{\epsilon}_s$, $U_p=\epsilon_p-\bar{\epsilon}_p$ in the orthonormal local orbital $\mathrm{s}\,|lj\alpha\rangle$; j denotes the two atoms in the unit cell labeled l, α represents s or p symmetry, and $\bar{\epsilon}_s$ and $\bar{\epsilon}_p$ are the concentration-weighted average values of s and p silicon and germanium term value energies.

The one-particle alloy Green's function is defined as

$$G_{\mathrm{alloy}}(Z)=\frac{1}{Z-H_{\mathrm{alloy}}}\ . \tag{2}$$

We are after the configuration average of this Green's function, which, in effective medium theory, is replaced by an effective Green's function G,

$$G(Z)=\frac{1}{Z-\overline{H}-\Sigma(Z)}\ , \tag{3}$$

where Σ is the self-energy. In CPA, we can now write $\Sigma=\sum_l(\Sigma_l)$, with Σ_l being an 8×8 matrix in the basis $|lj\alpha\rangle$ having the form

$$\underline{\Sigma}=\begin{vmatrix}\underline{A} & \underline{0}\\ \underline{0} & \underline{A}\end{vmatrix}\ , \tag{4}$$

where

$$\underline{A}=\begin{vmatrix}\Sigma_s & 0 & 0 & 0\\ 0 & \Sigma_p & 0 & 0\\ 0 & 0 & \Sigma_p & 0\\ 0 & 0 & 0 & \Sigma_p\end{vmatrix}\ .$$

Here Σ_s and Σ_p are the s and p parts of the self-energy. The Σ_s and Σ_p are determined from the conditions that the average atomic t matrix with respect to the CPA Green's function G is zero. With our ansatz for Σ, the matrix equation $\langle t\rangle=0$ reduces to two coupled equations $\langle t_s\rangle=0$ and $\langle t_p\rangle=0$, where the average is the concentration-weighted average $\langle Q\rangle=xQ^{\mathrm{Si}}+yQ^{\mathrm{Ge}}$, and the t is defined as

$$t_\alpha^\beta=(U_\alpha^\beta-\Sigma_\alpha)[1-F_\alpha(U_\alpha^\beta-\Sigma_\alpha)]^{-1}\ ,$$
$$(\alpha=s\text{ or }p\ ,\ \beta=\mathrm{Si}\text{ or }\mathrm{Ge})\ . \tag{5}$$

In the above expression F_a is the diagonal matrix element of G in the local basis

$$F_a(Z) \equiv \langle lja \,|\, G(Z) \,|\, lja \rangle \ .$$

Σ_s and Σ_p are coupled because F_s and F_p each contain both Σ_s and Σ_p.

An iterative average-t-matrix (IATA) procedure[19] is employed to solve the CPA equation. This procedure improves Σ_a upon a guessed solution Σ_a^0 through the following equations:

$$\begin{aligned} \Sigma_s &= \Sigma_s^0 + \langle t_s^0 \rangle (1 + F_s^0 \langle t_s^0 \rangle)^{-1} \ , \\ \Sigma_p &= \Sigma_p^0 + \langle t_p^0 \rangle (1 + F_p^0 \langle t_p^0 \rangle)^{-1} \ , \end{aligned} \tag{6}$$

where $\langle t_a^0 \rangle$ and F_a^0 are similar to those in Eq. (5) except that Σ_a^0 now replaces Σ_a. The most time-consuming calculation is then the computation of the local Green's functions F_s^0 and F_p^0, given by the BZ summation, e.g.,

$$F_s^0(Z) = \frac{1}{N} \sum_{\mathbf{k}} \left[\frac{1}{Z - \bar{H}(\mathbf{k}) - \Sigma_s^0} \right]_{11} ,$$

where the inverse of an 8×8 matrix is involved for every $\mathbf{k}$. This can be simplified by observing that Σ_a^0 has the same form as Σ in Eqs. (4) and (5) and that the 4×4 $\underline{A}$ matrix can be written as $\underline{A} = \Sigma_p \underline{1} + (\Sigma_s - \Sigma_p)\underline{J}$, where $\underline{1}$ is the identity matrix and

$$\underline{J} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix} .$$

Defining the matrix

$$\underline{Q} = (\Sigma_s^0 - \Sigma_p^0) \begin{vmatrix} \underline{J} & \underline{0} \\ \underline{0} & \underline{J} \end{vmatrix} ,$$

$F_s^0(Z)$ and $F_p^0(Z)$ can now be calculated from

$$\begin{aligned} F_s^0(Z) &= \frac{1}{N} \sum_{\mathbf{k}} g_{11}(\mathbf{k}, Z) \ , \\ F_p^0(Z) &= \frac{1}{3N} \sum_{\mathbf{k}} [g_{22}(\mathbf{k}, Z) + g_{33}(\mathbf{k}, Z) + g_{44}(\mathbf{k}, Z)] \ , \end{aligned} \tag{7}$$

where

$$g = g^0 + g^0 (1 - \sigma g^0)^{-1} \ , \tag{8}$$

with

$$g_{\alpha\beta}^0(\mathbf{k}, Z) = \sum_n \frac{U_{n\alpha}^*(\mathbf{k}) U_{n\beta}(\mathbf{k})}{Z - \epsilon_n(\mathbf{k}) - \Sigma_p^0} \ . \tag{9}$$

In Eq. (9), $\epsilon_n(\mathbf{k})$ is the band energy in VCA and $\{U_{\alpha n}(\mathbf{k})\}$ satisfy the following Eigen equation:

$$\sum_\beta \bar{H}_{\alpha\beta}(\mathbf{k}) U_{\beta n}(\mathbf{k}) = \epsilon_n(\mathbf{k}) U_{\alpha n}(\mathbf{k}) \ .$$

Because the σ matrix has only two nonzero elements, the matrix inversion in Eq. (8) is obtained analytically.

A substantial reduction in computer time is made possible by using an analytical continuation method.[32] In this method, Σ_s and Σ_p are calculated as a function of complex Z, and then, using the analytical properties of the self-energy and Green's functions, they are interpolated for real Z. Because the functions Σ_s, Σ_p, and G are smooth for complex Z, the CPA iterations and BZ integrations can be carried out with substantially less computer time.

For the concentrations $x = 0.10$ and 0.50, the L and X (Δ) gaps, respectively, are preferred. The L to X (Δ) crossover takes place near $x \approx 0.15$. The CPA correction to L and X edges at $x = 0.10$, 0.15, and 0.50 should be good enough to study the quantitative variation of band gap in Si_xGe_{1-x} alloys. Hence, the calculations are carried out for these three cases. In addition, because the experimental results are available for $x = 0.109$, CPA calculations are also done here for comparison. As expected, Σ_s is much larger than Σ_p for all the cases. The self-energy as a function of energy is plotted in Fig. 2 for an $x = 0.50$ alloy.

C. Off-diagonal CPA

As mentioned earlier, silicon and germanium differ in their lattice constant by $\sim 4\%$. In order to include the effect of the structural disorder, the CPA calculation is repeated next with off-diagonal (OD) disorder included. By an application of the molecular coherent-potential approximation (MCPA),[33] Hass *et al.* included OD disorder in the CPA calculation of $A'_xA''_{1-x}B$ semiconductor alloys.[34] Assuming that B atoms occupy the sites of an ordered zinc-blende virtual lattice, they modeled the dominant structural effect as the difference in A'-B and A''-B hopping matrix elements. Hence, the chemical and structural disorder effects are treated as random variations of $\epsilon^A, V_1^A, V_2^{AB}$, where the symbols have their usual meaning.[35]

The extension of the method to Si_xGe_{1-x} alloys is not straightforward, mainly because silicon and germanium can occupy both sublattices; hence, there can be no ordered virtual lattice in this case. If we choose the tetrahedral unit cell as the molecular unit for MCPA, we see that the disorder is not cell diagonal. However, by choosing an appropriate basis set, we can make the intercell interaction be the highest-order effect. We start with a hybrid basis $|lh\rangle_i$ obtained from the sp^3 hybrid orbitals.[35] The hybrids 1 through 4 ($i = 1$–4) are obtained from orbitals centered on a sublattice I site, and the states 5 through 8 ($i = 5$–8) are those from the orbitals located at the four nearest-neighbor sites on sublattice II. The Bloch basis states, corresponding to A_1, T_2 symmetries, $|\mathbf{k}\rangle_i$ located on a I site ($i = 1$–4) and II site ($i = 5$–8), are obtained from the corresponding hybrid states given by the relation

$$|l\rangle_i = \sum_{j=1}^{8} C_{ij} \, |lh\rangle_j \ , \tag{10}$$

where

$$\underline{C} = \begin{vmatrix} \underline{c}_1 & \underline{0} \\ \underline{0} & \underline{c}_1 \end{vmatrix}$$

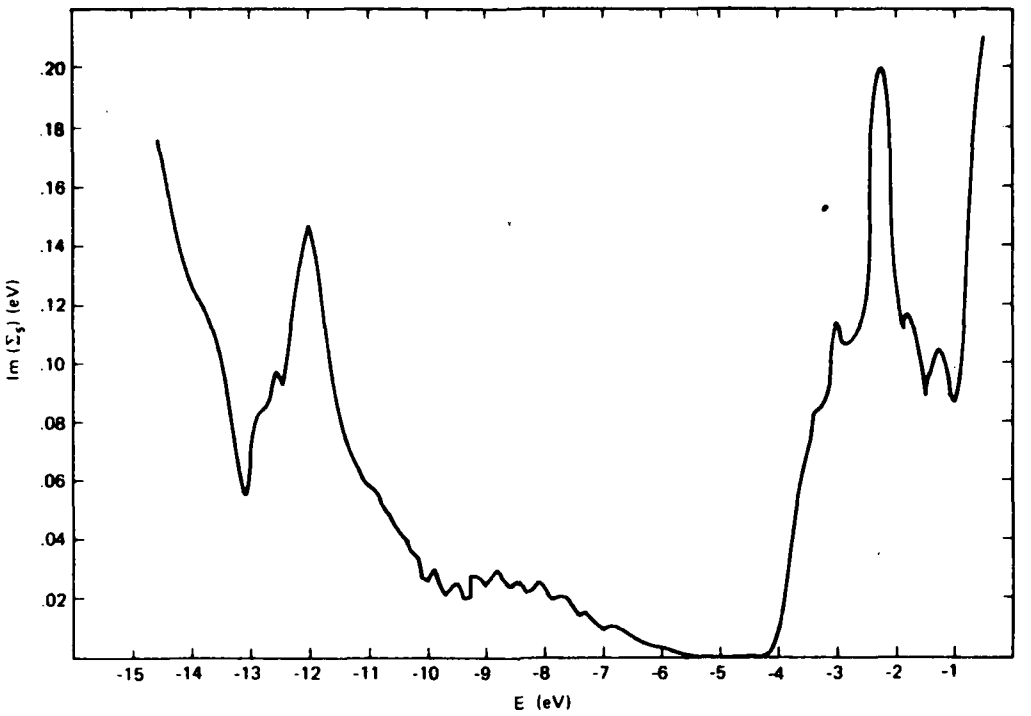


FIG. 2. Variation of the imaginary part of the self-energy Σ_s(CPA) as a function of energy for $x=0.50$.

and

$$\mathcal{L}_1=\frac{1}{2}\begin{vmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{vmatrix}.$$

An explicit definition of these orbitals can be found in Ref. 20. In this new basis, the self-energy Σ at the given site takes the form

$$\Sigma=\begin{vmatrix}Q_0&Q_2\\Q_2&Q_0'\end{vmatrix}, \tag{11}$$

where

$$Q_0=\begin{vmatrix}\Sigma_s&0&0&0\\0&\Sigma_p&0&0\\0&0&\Sigma_p&0\\0&0&0&\Sigma_p\end{vmatrix},$$

$$Q_2=\begin{vmatrix}\Sigma_s'&0&0&0\\0&\Sigma_p'&0&0\\0&0&\Sigma_p'&0\\0&0&0&\Sigma_p'\end{vmatrix}, \tag{12}$$

$$Q_0'=\begin{vmatrix}\Sigma_h'&0&0&0\\0&\Sigma_h'&0&0\\0&0&\Sigma_h'&0\\0&0&0&\Sigma_h'\end{vmatrix},$$

and

$$\Sigma_h'=\tfrac{1}{4}(3\Sigma_p+\Sigma_s), \tag{13}$$

where $(\Sigma_s,\Sigma_p),(\Sigma_s',\Sigma_p')$ are, respectively, the self-energies associated with diagonal and off-diagonal disorder corresponding s and p symmetries.

The self-energies can be obtained again from the IATA iteration procedure:

$$\Sigma=\Sigma_0+\langle\langle T\rangle\rangle(1+F\langle\langle T\rangle\rangle)^{-1}, \tag{14}$$

where

$$F_{ij}=\frac{1}{N}\sum_n\sum_{\mathbf{k}}\frac{Q_{in}^{\dagger}Q_{nj}}{Z-E_n(Z,\mathbf{k})}, \tag{15}$$

with

$$[\bar{H}(\mathbf{k})+\Sigma]Q=E(Z,\mathbf{k})Q,$$

$$\langle\langle T\rangle\rangle=x\langle T_A\rangle+y\langle T_B\rangle,$$

and

$$\langle T_A\rangle=x^4t_{A_4}^A+4x^3yt_{A_3B}^A+6x^2y^2t_{A_2B_2}^A+4y^3xt_{AB_3}^A+y^4t_{B_4}^A, \quad A\equiv\text{Si} \tag{16}$$

with a similar expression for $\langle T_B\rangle$. Physically, for a given A atom at the center, the other four atoms in the molecular unit cell can be all A atoms, three A atoms and one B atom, two of each, one A atom and three B atoms, or all four can be B atoms. $\langle T_A\rangle$ or $\langle T_B\rangle$ represents the configuration-averaged t matrices, and $\langle\langle T\rangle\rangle$ is the concentration-weighted average of the configuration. By exploiting the symmetry, as seen in Eq. (11), one can

reduce this problem to solving two 2×2 coupled matrix equations. Equation (14) can be iterated to obtain Σ_s, Σ_p, Σ_s', and Σ_p'. After every iteration, we get a new set of Σ_s, Σ_p, Σ_s', Σ_p', and Σ_h'. The new set has not been tested to see if Σ_h' is still given by Eq. (13). In our calculation, we did not iterate to obtain a new Σ_h'; the error introduced by this approximation is expected to be very small. As in the case of diagonal CPA, the computation can be substantially reduced by the method of analytical continuation.[32]

IV. DISCUSSION

A. E_0 and E_1 optical transitions

The VCA values of E_0 (Γ_{2c}'-Γ_{25v}) and E_0' (Γ_{15c}-Γ_{25v}) and their measured values are plotted as a function of x in Fig. 1(e). Because the measurements[36] are made at room temperature, the experimental values are smaller than the values calculated from the zero-temperature band structure. Inclusion of the relativistic effects, which are not present in our calculations, is expected to form a more accurate basis for comparison with the experiments. As seen from Fig. 1(e), the theoretical and the experimental values both have a linear variation with x. Similar calculations of E_1 ($L_{1c}-L_{3v}'$) also have a linear variation on x and are in qualitative agreement with experiments.[36]

From the CPA self-energies Σ_s and Σ_p, it is straightforward to calculate the correction to the VCA bands. The calculated complex band structure is plotted for $x=0.50$ in Fig. 3. The CPA corrections are shown only in the vicinity of the band gap. The shaded portion represents the half-width of that energy state. Because s scattering is dominant in these alloys, we see that the major disorder lies in the conduction band. The topmost valence band, with its rich p content, is least affected. The CPA band structure is used to calculate the E_0 and E_1 peak positions for $x=0.10$, 0.109, 0.15, and 0.50 concentrations. The calculations and the data from Ref. 36 show a small bowing that is not seen on the scale of Fig. 1(e).

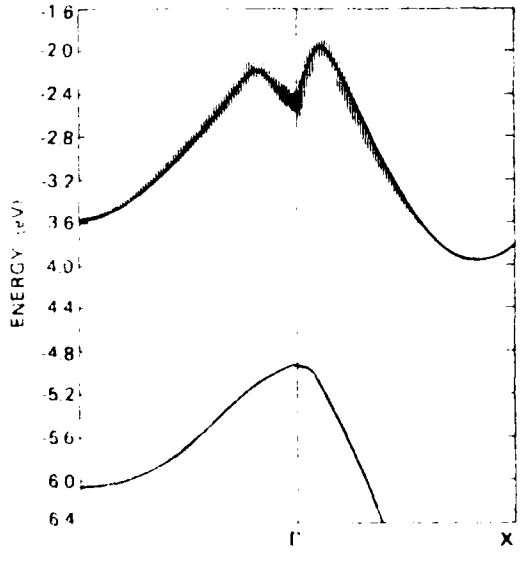


FIG. 3. Calculated CPA complex band structure of the $Si_{0.5}Ge_{0.5}$ alloy. Only the bands in the vicinity of the energy gap are shown. The shaded portion represents the alloy broadening

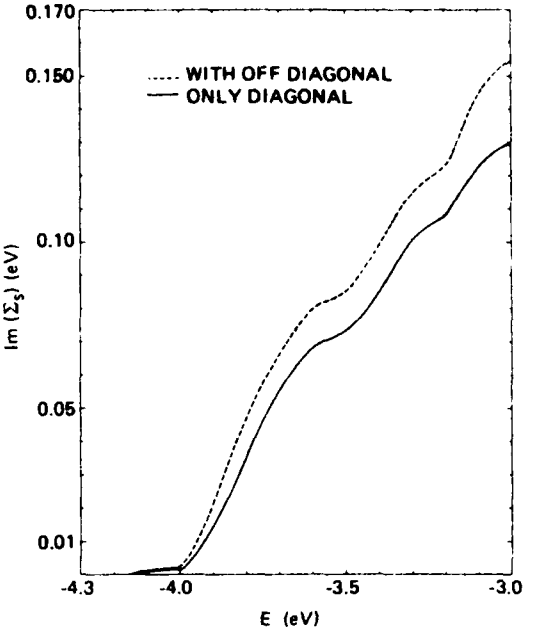


FIG. 4. Imaginary part of Σ_s(CPA) and Σ_s(MCPA) as a function of energy for $x=0.50$ alloy.

The self-energies Σ_s, Σ_p, Σ_p', and Σ_s' are calculated in MCPA for the $x=0.50$ alloy. As in the case of CPA, the self-energies associated with s symmetry are much larger than the ones associated with the p symmetry. While Σ_s' is found to be very small, Σ_p' is at least an order of magnitude smaller—almost zero. However, the $\mathrm{Im}\Sigma_s$ obtained by CPA and MCPA differ considerably. As seen from Fig. 4, the difference increases as one goes away from the band edge. Therefore, the lifetime associated with the alloy disorder is decreased by the inclusion of OD structural disorder. In addition, the OD disorder lowers the conduction band, introducing an extra bowing. The E_0 and E_1

TABLE III. Calculated values of E_0,E_1 and their respective half-widths $\Delta(E_0)$ and $\Delta(E_1)$ (all energies are in eV).

X	Quantity	VCA	CPA	MCPA
0.10	E_0	1.290	1.248	
	$\Delta(E_0)$		0.011	
	E_1	2.016	1.995	
	$\Delta(E_0)$		0.001	
0.109	E_0	2.028	2.009	
	$\Delta(E_0)$		0.013	
	E_1	1.318	1.275	
	$\Delta(E_1)$		0.002	
0.15	E_0	1.442	1.382	
	$\Delta(E_0)$		0.032	
	E_1	2.083	2.051	
	$\Delta(E_1)$		0.002	
0.50	E_0	2.517	2.418	2.391
	$\Delta(E_0)$		0.186	0.206
	E_1	2.578	2.510	2.498
	$\Delta(E_1)$		0.0308	0.0319

values are reduced by 27 and 12 meV, respectively. The VCA, CPA, and MCPA values of E_0 and E_1 are listed in Table III.

The half-width of the alloy states is calculated from the imaginary part of the CPA self-energies. The half-width of the lowest-lying conduction band of $Si_{0.5}Ge_{0.5}$ alloy is plotted in Fig. 5 as a function of K_x in the [100] direction. The calculated half-width is 186 meV for the Γ_2' state and decreases to zero at the band edge. Because of the negligible alloy broadening of the topmost valence-band state, the half-width corresponding to the E_0 transition, $\Delta(E_0)$, is 186 meV, which is approximately one-half of the previously published CPA results.[13] The CPA value of the half-width corresponding to the E_1 transition, $\Delta(E_1)$, is 31 meV. Because of the increase in the imaginary part of the self-energies, the MCPA values of the half-widths of the E_0 and E_1 transitions are 206 and 32 meV, respectively. Because the complete E_0 peak is not shown in the published electroreflectance spectrum,[36] it is difficult to estimate the corresponding half-width. However, one can conclude from the spectrum of the $x=0.458$ alloy that the half-width of the E_1 transition is considerably smaller ($\simeq$50 meV) than that of the E_0 transition. The agreement between the experimental and the theoretical values can be regarded as good because there are errors in estimating the width from the published spectra, and we have neglected the extrinsic broadening due to the apparatus used in the experiments.

In order to make a more accurate comparison with the experiments, the CPA values of $\Delta(E_0)$ and $\Delta(E_1)$ are calculated for the $x=0.109$ alloy. The calculated half-widths of the E_0 and E_1 transitions are 13 and 2 meV, respectively. From the spectrum, we estimate the corresponding values to be 8–15 and 3–6 meV. We see that CPA values are in excellent agreement with these experiments. Because x is small, the inclusion of off-diagonal disorder is not expected to change the calculated values significantly.

B. Energy gap

The fundamental gaps of these alloys are calculated as a function of concentration. The VCA gap is an increasing function of x with a slope discontinuity at $x\simeq 0.11$. The conduction-band minimum changes from L point to X (Δ) point at this crossover. In addition to the band gap, the effective electron masses and the band edge K_0 are also calculated. When X (Δ) gap is preferred, the band edge moves linearly from $\mathbf{k}$ at $(0.9,0,0)_{x=0.15}$ to $(0.8,0,0)_{x=1}$. The effective masses at a given minimum increase linearly from their pure germanium values to the corresponding pure silicon values.

Using CPA self-energies, the band gap, band masses, and the band edge are also calculated. The position of the band minimum did not change by virtue of the inclusion of off-diagonal disorder. While the effective transverse mass remains almost the same as the VCA value, the longitudinal mass has a maximum of 12% enhancement. Because the real part of CPA self-energies is negative in the forbidden gap region, an extra bowing is introduced to the VCA energy gap. Because of this bowing, the L-X (Δ) crossover takes place near $x\simeq 0.13$. The VCA, CPA, and experimental[36] bowing parameters are 0.06, 0.18, and 0.24, respectively. The calculated energy gap is plotted as a function of x in Fig. 6.

Because of the negligible change in the effective masses, the corresponding values in the pure materials are used in the calculation of the alloy-scattering-limited electron mobility. The CPA X-gap E_g^X and L-gap E_g^L are fitted to a polynomial form. The generalized Brooks's formula that is applicable to the alloys with an indirect gap and multiple bands is used.[15] The calculated electron-drift mobility and the experimental Hall mobility[1] are plotted in Fig. 7, where the theory explains the qualitative behavior of experimental results.[1] As observed,[5] even a few atomic percent alloy concentration can reduce the drift mobility substantially. It can be seen that the rate of decrease near $x=0$ and $x=1$ are quite different. This is because the L edge has more s content than the X edge. Because the s

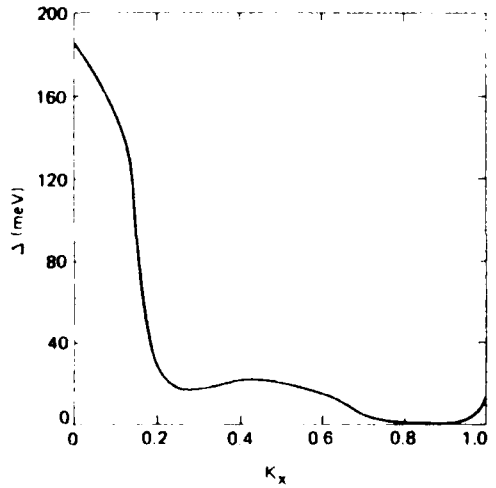


FIG. 5. Variation in the width of the lowest-lying conduction band as a function of K_x in the [100] direction for the $x=0.50$ alloy.

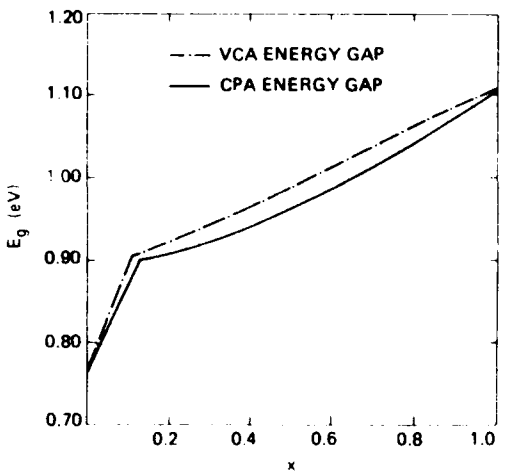


FIG. 6. Variation of the VCA energy gap (dash-dotted line) and the CPA energy gap (solid line) as a function of x.

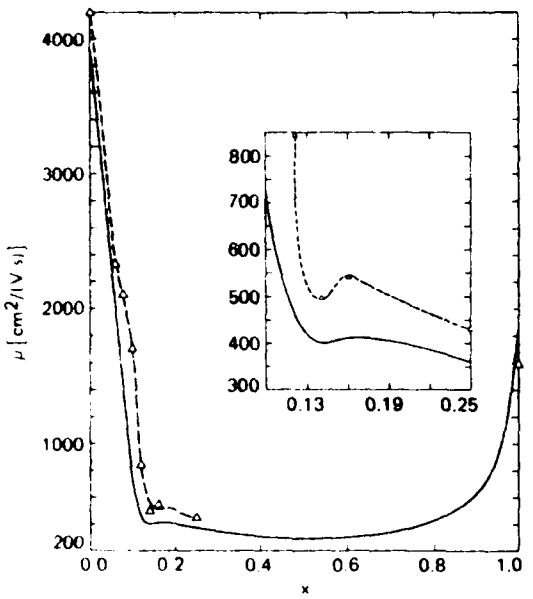


FIG. 7. Calculated drift mobility (solid line) and the experimental (dashed line) Hall mobility (Ref. 1) as a function of x.

scattering is dominant in these alloys, the L electrons are scattered more than the X electrons. Precisely for this reason, one observes a dip in the mobility near the L to X (Δ) crossover. For $x < 0.13$, the minimum gap is the L gap. After the crossover, the minimum gap is the X (Δ) gap, and the reduced alloy scattering increases the average mobility. For still larger x, the mobility decreases because of the increased alloy disorder. All these features are clearly seen in Fig. 7. While our calculations include the intervalley scattering mediated by alloy disorder, the effect of other scattering mechanisms is expected to increase the dip near the crossover.

The calculated alloy scattering rate for the holes is several orders smaller than that for electrons, because (1) the valence-band edge has dominant p content, (2) the p-scattering parameter ($\Delta\epsilon_p = 0.21$) is only $\frac{1}{7}$ of $\Delta\epsilon_s$, which alone decreases the scattering rate for holes by a factor of 50, and (3) finally, the imaginary part of the self-energy is proportional to the density of states, which approaches zero at the band edge. Hence, the hole mobility in this system is insensitive to alloy disorder.

In MCPA, the conduction band is pushed down, because of an increase in the imaginary part of the self-energy, giving rise to an additional bowing in the fundamental gap. For an $x = 0.50$ alloy, the gap is reduced by 7 meV. The bowing parameter, including the MCPA correction, is 0.21, which is in excellent agreement with experiment.[37]

It is interesting to compare the results of our calculations with those of Hass *et al.*[34] In their calculations on the $Ga_{1-x}In_xAs$ alloy, CPA introduced an extra bowing in the fundamental gap. However, after the MCPA corrections, the total scattering was diminished and the results were similar to VCA results. These results were explained in terms of the relative strength and sign of the atomic term values and V_2^{AB}. We extend their argument to Si_xGe_{1-x} alloys. The hybrid level of silicon is higher than that of germanium. Because of its shorter bond length, the V_2 of silicon is larger than that of germanium. Thus, in this case, both effects combine to give more disorder in the conduction and valence bands. Therefore, the scattering is enhanced in these alloys. This explains the increase in the imaginary part of the self-energy due to inclusion of OD disorder in our calculation.

In conclusion, we have incorporated both chemical and structural disorder into the calculation of the CPA band structure of Si_xGe_{1-x} alloys. The calculation, based on a realistic band structure of silicon and germanium, suggests that the band gap is an increasing function of x with a slope discontinuity at $x \simeq 0.13$. The linewidths of the E_0 and E_1 transitions calculated by CPA and MCPA are in good agreement with experiments. The effects of structural disorder on diagonal CPA for the Ge-Si alloy system will be difficult to test experimentally. MCPA decreases the band gap only slightly. However, unlike the GaInAs case, the imaginary part of the s self-energy increases with increasing energy and serves as a warning that all systems will not exhibit the same behavior. Therefore, in Ge-Si alloys, the place to look for experimentally significant differences between CPA and MCPA predictions is in the high-lying transitions. However, even in the E_0 transition, the MCPA linewidth at $x = 0.5$ is still only 10% larger than the corresponding CPA value. The calculated alloy-scattering-limited electron-drift mobility is in qualitative agreement with the observed Hall mobilities.

ACKNOWLEDGMENTS

This work was supported in part by U.S. Air Force Office of Scientific Research Grant No. AFOSR-84-0282 and Defense Advanced Research Projects Agency (DARPA)/AFOSR Project No. PR:FQ8671-851100. One of us (A.-B.C.) would like to thank Professor W. E. Spicer for his hospitality at Stanford University.

[1] M. Glicksman, Phys. Rev. 111, 125 (1958); 100, 1146 (1955).
[2] B. A. Bunker, S. L. Hulbert, J. P. Stott, and F. C. Brown, Phys. Rev. Lett. 53, 2157 (1984).
[3] H. J. Lee, L. Y. Juravel, and J. C. Woolley, Phys. Rev. B 21, 659 (1980).
[4] N. Lifshitz, A. Jayaraman, and R. A. Logan, Phys. Rev. B 21, 670 (1980).
[5] H. M. Manasevit, I. S. Gergis, and A. B. Jones, Appl. Phys. Lett. 41, 464 (1982); J. Electron. Mater. 12, 637 (1983).
[6] S. Krishnamurthy and J. A. Moriarty, Superlat. Microstruct. 1, 209 (1985); Phys. Rev. B 32, 1027 (1985).
[7] J. C. Bean, L. C. Feldman, A. T. Flory, S. Nakahara, and I. K. Robinson, J. Vac. Sci. Technol. A 2, 436 (1984).
[8] R. People, J. C. Bean, D. V. Lang, A. M. Sergent, H. L. Storm-

er, K. W. Wecht, R. T. Lynch, and K. Baldwin, Appl. Phys. Lett. **45**, 1231 (1984).
[9]F. Cerdeira, A. Pinczuk, and J. C. Bean, Phys. Rev. B **31**, 1202 (1985).
[10]T. P. Pearsall, F. H. Pollak, and J. C. Bean, Bull. Am. Phys. Soc. **30**, 266 (1985).
[11]K. E. Newman and J. D. Dow, Phys. Rev. B **30**, 1929 (1984).
[12]M. Z. Huang and W. Y. Ching, Superlat. Microstruct. **1**, 137 (1985).
[13]D. Stroud and H. Ehrenreich, Phys. Rev. B **2**, 3197 (1970).
[14]S. Krishnamurthy, A. Sher, and A.-B. Chen, Phys. Rev. Lett. **55**, 320 (1985).
[15]S. Krishnamurthy, A. Sher, and A.-B. Chen, Appl. Phys. Lett. **47**, 160 (1985).
[16]A.-B. Chen and A. Sher, Phys. Rev. B **22**, 3886 (1980).
[17]E. O. Kane, Phys. Rev. B **13**, 3478 (1976).
[18]D. J. Chadi, Phys. Rev. B **16**, 3572 (1977).
[19]A.-B. Chen and A. Sher, Phys. Rev. B **23**, 5360 (1981).
[20]A.-B. Chen and A. Sher, Phys. Rev. B **26**, 6603 (1982).
[21]R. R. L. Zucca, J. P. Walter, Y. R. Shen, and M. L. Cohen, Solid State Commun. **8**, 627 (1970).
[22]R. R. L. Zucca and Y. R. Shen, Phys. Rev. B **1**, 2668 (1970).
[23]M. Welkowsky and R. Braunstein, Phys. Rev. B **5**, 497 (1972).
[24]R. A. Pollak, L. Ley, S. Kowalcyzk, D. A. Shirley, J. Joannopolos, D. J. Chadi, and M. L. Cohen, Phys. Rev. Lett. **29**, 1103 (1973).
[25]L. Ley, R. A. Pollak, F. R. McFeely, S. P. Kowalcyzk, and D. A. Shirley, Phys. Rev. B **9**, 600 (1974).
[26]W. D. Grobman and D. E. Eastman, Phys. Rev. Lett. **29**, 1508 (1972).
[27]D. E. Eastman, W. D. Grobman, J. L. Freeouf, and M. Erbudak, Phys. Rev. B **9**, 3473 (1974).
[28]W. E. Spicer and R. C. Eden, *Proceedings of the Ninth International Conference on the Physics of Semiconductors, 1968* (Nauka, Leningrad, 1969), Vol. 1, p. 61.
[29]G. W. Gobeli and F. G. Allen, Phys. Rev. **245**, A137 (1965).
[30]J. R. Cheliskowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976), and references therein.
[31]J. A. Moriarty and S. Krishnamurthy, J. Appl. Phys. **54**, 1892 (1983).
[32]K. C. Hass, B. Velicky, and H. Ehrenreich, Phys. Rev. B **29**, 3697 (1984).
[33]F. Ducastelle, J. Phys. C **7**, 1795 (1974).
[34]K. C. Hass, R. J. Lampert, and H. Ehrenreich, Phys. Rev. Lett. **52**, 77 (1984).
[35]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), p. 65.
[36]J. S. Kline, F. H. Pollak, and M. Cardona, Helv. Phys. Acta **41**, 968 (1968).
[37]R. Braunstein, A. R. Moore, and F. Herman, Phys. Rev. **109**, 695 (1958).

END

10-86

DTIC